C-27

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXI, Number 2
APRIL, 1939

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

THE AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. COHEN at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

SOME RESULTS IN THE ARITHMETIC THEORY OF ALGEBRAIC VARIETIES.*

By Oscar Zariski.

Introduction. In Part I we treat systematically some basic questions of the theory of singularities of an algebraic variety. The main results bear upon various characterizations of simple points of a variety. We begin with an ideal theoretic definition of a simple point (section 2). This definition is given of terms of relative unramified prime zero-dimensional ideals. We next characterize a simple point by the existence of a local uniformization satisfying certain conditions (sections 3 and 4). At the same time we obtain a characterization of a simple point by an intrinsic (non-relative) property of the corresponding prime zero-dimensional ideal (Theorem 3. 2). Finally we connect the singular points of a variety with the properties of the different (section (5-8)) and we exhibit an ideal whose variety is the manifold of singular points of the given variety (sections 8-9).

In Part II we derive some properties of the conductor of a finite integral domain $\mathfrak{o}$ with respect to its integral closure $\mathfrak{o}^*$. Those properties concern the usual question of the decomposition in $\mathfrak{o}^*$ of a prime ideal in $\mathfrak{o}$. The result is complete in the case that $\mathfrak{o}$ is generated by the independent variables and a primitive element (sections 11-12). Also here it was possible to obtain an intrinsic result valid in $\mathfrak{o}^*$ (section 13).

The contents of Parts I and II are, in the main, generalizations of well-known theorems in the arithmetic theory of algebraic functions of one variable.

The contributions of Parts III and IV are new also in the case of functions of one variable. Here we introduce the concept of a *normal* variety, both in the affine and in the projective space, and we are led to a geometric interpretation of the operation of integral closure. The importance of normal varieties is due to their following two properties: 1) *the singular manifold of a normal V_r is of dimension* $\leqq r-2$ (in particular a normal curve (V_1) is free from singularities); 2) *the system of hyperplane sections of a normal V_r is complete.* There is a definite class of normal varieties associated with and birationally equivalent to a given variety V_r. This class is obtained by a process of integral closure carried out in a suitable fashion for varieties in projective spaces. It turns out that the varieties of this class are those on

---

* Received October 7, 1938.

which the hyperplanes cut out the complete systems $|hC|$, sufficiently high multiples of the system $|C|$ of hyperplane sections of the given V_r. These results seem to point to a fruitful arithmetic approach to questions of the birational theory of varieties.

The special birational transformations effected by the operation of integral closure, and the properties of normal surfaces, play an essential rôle in our arithmetic proof for the reduction of singularities of an algebraic surface. This proof will be published in the July issue of the *Annals of Mathematics.*

Although the underlying field of coefficients is supposed throughout to be of characteristic zero, the proofs remain valid for any characteristic, provided only separable extensions are being considered. The separability is no restriction from the birational point of view, or even from the projective point of view.

Many of the results of Parts I, II, and III have been announced by the author without proof in a Note of the *Proceedings of the National Academy of Sciences* [10].

I. Simple and multiple points of an algebraic variety.

1. We consider an algebraic irreducible r-dimensional variety V_r, in an affine space $S_n(x_1, \cdots, x_n)$, over an algebraically closed ground field K of characteristic zero. Let $\xi_1, \cdots, \xi_n$ be the coördinates of the general point of V_r and let $\Sigma = \mathsf{K}(\xi_1, \cdots, \xi_n)$ be the field of rational functions on V_r, of degree of transcendency r over K. We denote by $\mathfrak{o}$ the ring $\mathsf{K}[\xi_1, \cdots, \xi_n]$, whose elements are polynomials in the ξ's. The defining ideal of V_r in the polynomial ring $\mathsf{K}[x] (= \mathsf{K}[x_1, \cdots, x_n])$ is the prime r-dimensional ideal $\mathfrak{p}'_r$, consisting of all polynomials $f(x_1, \cdots, x_n)$ such that $f(\xi_1, \cdots, \xi_n) = 0$; $\mathfrak{o}$ is the ring of residual classes $\mathsf{K}[x]/\mathfrak{p}'_r$ and Σ is the quotient field of $\mathfrak{o}$.

The polynomials in $\mathsf{K}[x]$ which vanish at a given point $P(a_1, \cdots, a_n)$ of S_n form a prime 0-dimensional ideal $\mathfrak{p}'_0 = (x_1 - a_1, \cdots, x_n - a_n)$. The point P is on V_r if and only if $\mathfrak{p}'_r \subset \mathfrak{p}'_0$. The homomorphism between $\mathsf{K}[x]$ and $\mathsf{K}[\xi]$ sets up a one-to-one correspondence between the prime 0-dimensional ideals $\mathfrak{p}'_0$ in $\mathsf{K}[x]$ which contain $\mathfrak{p}'_r$ and the prime 0-dimensional ideals $\mathfrak{p}_0$ in $\mathfrak{o}$ $(= \mathsf{K}[\xi])$; here $\mathfrak{p}_0 = \mathfrak{p}'_0/\mathfrak{p}'_r$. Thus there is a one-to-one correspondence betwen the points of V_r (in the affine S_n) and the prime 0-dimensional ideals in $\mathfrak{o}$: if $P(a_1, \cdots, a_n)$ is a point on V_r, then the ideal $\mathfrak{p}_0 = (\xi_1 - a_1, \cdots, \xi_n - a_n)$ in $\mathfrak{o}$ is prime and 0-dimensional, and conversely. If P is not on V_r, then $(\xi_1 - a_1, \cdots, \xi_n - a_n)$ is the unit ideal.

If $\mathfrak{p}_0$ is any prime 0-dimensional ideal in $\mathfrak{o}$, then the ring $\mathfrak{o}/\mathfrak{p}_0$ contains a field $\mathsf{K}^* \cong \mathsf{K}$ (if $c_1, c_2 \subset \mathsf{K}$ and $c_1 \neq c_2$, then $c_1 - c_2 \not\equiv 0(\mathfrak{p}_0)$, since $\mathfrak{p}_0$ is

not the unit ideal), and moreover every element of $\mathfrak{o}/\mathfrak{p}_0$ is algebraic over K^*. Since K^* is assumed algebraically closed, it follows that $\mathfrak{o}/\mathfrak{p}_0$ coincides with the field K^*.[1] Thus every element ω in $\mathfrak{o}$ satisfies a congruence of the form $\omega \equiv c(\mathfrak{p}_0)$. We shall say that ω *has the value* c *at* P. In particular, $\xi_i \equiv a_i(\mathfrak{p}_0)$ and to $\mathfrak{p}_0$ there corresponds the point $P(a_1, \cdots, a_n)$ on V_r.

2. DEFINITION. *A point* $P(a_1, \cdots, a_n)$ *on* V_r *is said to be a simple point of* V_r, *if there exist, in* $\mathfrak{o}$, r *elements* $\eta_1, \cdots, \eta_r$ *such that the ideal* $\mathfrak{A} = \mathfrak{o} \cdot (\eta_1, \cdots, \eta_r)$ *is divisible by (is contained in)* $\mathfrak{p}_0$ $(= \mathfrak{o} \cdot (\xi_1 - a_1, \cdots, \xi_n - a_n))$ *but is not divisible by any proper primary ideal belonging to* $\mathfrak{p}_0$.

The condition in the above definition is equivalent to the following: $\mathfrak{p}_0$ *must be an isolated component of the ideal* $\mathfrak{A}$. In fact, $\mathfrak{A} \equiv 0(\mathfrak{p}_0)$ and if $\mathfrak{p}_0$ were not an isolated component of $\mathfrak{A}$, then $\mathfrak{A}$ must be divisible by some prime ideal $\mathfrak{p}$ of dimension > 0 and contained in $\mathfrak{p}_0$ (since, by hypothesis, no primary ideal belonging to $\mathfrak{p}_0$ can be a component of $\mathfrak{A}$). This, however, is impossible, since it can be easily shown that any such prime ideal $\mathfrak{p}$ is also a multiple of some primary ideal belonging to $\mathfrak{p}_0$.[2]

We proceed to derive properties of a simple point which we shall have occasion to use in the sequel and which will also bear upon the geometric content of our definition.

Let $P(a_1, \cdots, a_n)$ be a simple point of V_r and let, according to our definition of a simple point, $\eta_1, \cdots, \eta_r$ be r elements in $\mathfrak{o}$ such that the ideal $\mathfrak{p}_0 = (\xi_1 - a_1, \cdots, \xi_n - a_n)$ is an isolated component of the ideal $\mathfrak{A} = (\eta_1, \cdots, \eta_r)$. We may assume, without loss of generality, that P is at the origin of coördinates, whence $\mathfrak{p}_0 = (\xi_1, \cdots, \xi_n)$.

The elements η_i are polynomials in the ξ's. In each of these polynomials the constant term is zero, since $\eta_i \equiv 0(\mathfrak{p}_0)$. Let

$$\eta_i = c_{i1}\xi_1 + \cdots + c_{in}\xi_n + \text{terms of degree} \geqq 2. \tag{1}$$

We assert that *the* r *linear forms* $\sum_{j=1}^{n} c_{ij}\xi_j$ *are linearly independent modulo* $\mathfrak{p}_0^2$. To see this, we observe that if this were not the case, then also the r elements η_i would be linearly dependent modulo $\mathfrak{p}_0^2$, since $\eta_i - \sum_{j=1}^{n} c_{ij}\xi_j \equiv 0(\mathfrak{p}_0^2)$. Let, say,

---

[1] That $\mathfrak{o}/\mathfrak{p}_0$ is a field follows also from the fact that $\mathfrak{p}_0$ has no divisors other than $\mathfrak{p}_0$ and the unit ideal.

[2] In the homomorphism $\mathfrak{o} \simeq \bar{\mathfrak{o}} = \mathfrak{o}/\mathfrak{p}$, the ideal $\mathfrak{p}_0$, as a divisor of $\mathfrak{p}$, is mapped upon a 0-dimensional ideal $\bar{\mathfrak{p}}_0$ in $\bar{\mathfrak{o}}$, and the primary ideals belonging to $\bar{\mathfrak{p}}_0$ correspond to primary ideals in $\mathfrak{o}$ belonging to $\mathfrak{p}_0$ and containing $\mathfrak{p}$.

$$d_1\eta_1 + \cdots + d_r\eta_r \equiv 0(\mathfrak{p}_0^2),$$

where $d_1, \cdots, d_r$ are in K and are not all zero. Let $d_r \neq 0$. It follows from the above congruence that

$$\eta_r \equiv 0(\eta_1, \cdots, \eta_{r-1}, \mathfrak{p}_0^2). \tag{2}$$

We consider the ideal $\mathfrak{B} = (\eta_1, \cdots, \eta_{r-1})$. We observe that $\mathfrak{B}$ is not the unit ideal, since $\mathfrak{B} \subseteq \mathfrak{A} \subseteq \mathfrak{p}_0$. Hence, by a well known theorem [3] the minimal (non-imbedded) prime ideals of $\mathfrak{B}$ are all of dimension not less than 1. Since $\mathfrak{B} \subseteq \mathfrak{p}_0$, $\mathfrak{p}_0$ must divide at least one of the minimal prime ideals of $\mathfrak{B}$, say $\mathfrak{p}'$. It follows, as has been pointed out in footnote 2, that there exist primary ideals belonging to $\mathfrak{p}_0$ (and distinct from $\mathfrak{p}_0$) which divide $\mathfrak{p}'$. There will also then exist a maximal primary ideal $\mathfrak{q}_0$ with this same property, i. e. a primary ideal $\mathfrak{q}_0$ belonging to $\mathfrak{p}_0$ and such that there are no ideals between $\mathfrak{p}_0$ and $\mathfrak{q}_0$. Now it is well known that each maximal primary ideal of $\mathfrak{p}_0$ is a divisor of $\mathfrak{p}_0^2$.[4] Hence $\mathfrak{p}_0^2 \equiv 0(\mathfrak{q}_0)$, and since also $\mathfrak{B} \equiv 0(\mathfrak{p}') \equiv 0(\mathfrak{q}_0)$, it follows, by (2), $\eta_r \equiv 0(\mathfrak{q}_0)$. Hence $\mathfrak{A} \equiv 0(\mathfrak{q}_0)$, in contradiction with the hypothesis that $\mathfrak{p}_0$ is an isolated component of $\mathfrak{A}$.

In view of the linear independence of the forms $\sum_{j=1}^{n} c_{ij}\xi_j \bmod \mathfrak{p}_0^2$, it follows, *a fortiori*, that the matrix (c_{ij}) is of rank r. Hence by means of a non-singular linear homogeneous transformation of the coördinates ξ_i of the general point of V_r, it can be arranged that the elements η_i have the following form:

$$\eta_i = \xi_i + f_i(\xi_1, \cdots, \xi_n), \qquad (i = 1, 2, \cdots, r), \tag{3}$$

where f_i is a polynomial whose terms are all of degree $\geqq 2$. Now let ω be any element in $\mathfrak{p}_0$. Since $\mathfrak{p}_0$ is an isolated component of $\mathfrak{A}$, it follows that there exist an element α in $\mathfrak{o}$, such that $\alpha \not\equiv 0(\mathfrak{p}_0)$ and $\alpha\omega \equiv 0(\mathfrak{A})$. Since $\alpha \equiv c(\mathfrak{p}_0)$, $c \neq 0$, $c \subset \mathsf{K}$, this implies that $\omega \equiv 0(\mathfrak{A}, \mathfrak{p}_0^2)$ i. e.

$$\omega = A_1\eta_1 + \cdots + A_r\eta_r + \beta, \qquad A_i \subset \mathfrak{o},\ \beta \equiv 0(\mathfrak{p}_0^2).$$

Replacing the η_i by their expressions in (3) and observing that β, as an ele-

[3] See [2], p. 43 and the references on p. 45 to Macaulay and van der Waerden.

[4] See, for instance, [1], theorem 2, p. 529. Although the assertion is there proved only for rings in which the "weak" Doppelkettensatz holds true, the proof of this theorem, as well as of Theorem 3 and Corollary on p. 529, *loc. cit.*, carry over to *zero-dimensional* ideals in arbitrary finite integral domains. On the other hand, our assertion is practically trivial if we observe that it obviously holds in the polynomial ring $\mathsf{K}[x_1, \cdots, x_n]$ for 0-dimensional ideals and if we consider the homomorphism between $\mathsf{K}[x]$ and $\mathsf{K}[\xi]$.

ment of $\mathfrak{p}_0{}^2$, can be expressed as a polynomial in the ξ's in which all the terms are of degree $\geqq 2$, we see that our arbitrary element ω in $\mathfrak{p}_0$ can be put in the form:

$$\omega = c_1\xi_1 + \cdots + c_r\xi_r + g(\xi_1, \cdots, \xi_n), \tag{4}$$

where g contains only terms of degree $\geqq 2$. Or, in other words: there exist constants $c_1, \cdots, c_r$ such that

$$\omega \equiv c_1\xi_1 + \cdots + c_r\xi_r(\mathfrak{p}_0{}^2). \tag{5}$$

THEOREM 1. *Let* $\omega_1, \omega_2, \cdots, \omega_r$ *be elements of* $\mathfrak{p}_0$; $\omega_i \equiv \sum_{j=1}^{r} c_{ij}\xi_j(\mathfrak{p}_0{}^2)$. *A necessary and sufficient condition that* $\mathfrak{p}_0$ *be an isolated component of the ideal* $(\omega_1, \cdots, \omega_r)$ *is that the determinant* $| c_{ij} |$ *be different from zero.*

Proof. Assume $| c_{ij} | \neq 0$. Given any element ω in $\mathfrak{p}_0$, it is then possible to find constants $d_1, \cdots, d_r$, such that $\omega \equiv d_1\omega_1 + \cdots + d_r\omega_r(\mathfrak{p}_0{}^2)$. Hence $\mathfrak{p}_0 = (\omega_1, \cdots, \omega_r, \mathfrak{p}_0{}^2)$, and this implies that $\mathfrak{p}_0$ is an isolated component of the ideal $(\omega_1, \cdots, \omega_r)$ (in view of the fact that $\mathfrak{p}_0{}^2$ is a multiple of every maximal primary ideal belonging to $\mathfrak{p}_0$; see footnote 4).

Assume $| c_{ij} | = 0$. There exist then constants $d_1, \cdots, d_r$, not all zero, such that $d_1\omega_1 + \cdots + d_r\omega_r \equiv 0(\mathfrak{p}_0{}^2)$. If, for instance, $d_r \neq 0$, then it follows that $\omega_r \equiv 0(\omega_1, \cdots, \omega_{r-1}, \mathfrak{p}_0{}^2)$. This congruence is analogous to the congruence (2), encountered above, and leads therefore to a similar conclusion, e. g. that the ideal $(\omega_1, \cdots, \omega_r)$ is a multiple of some maximal primary ideal $\mathfrak{q}_0$ belonging to $\mathfrak{p}_0$, q. e. d.

In particular, if we take for $\omega_1, \cdots, \omega_r$ linear forms $\bar{\xi}_1, \cdots, \bar{\xi}_r$ in $\xi_1, \cdots, \xi_n$, then, for non-special values of the coefficients of these forms, the determinant $| c_{ij} |$ will be different from zero. Moreover, by a well known "normalization theorem,"[5] for non-special values of these coefficients the elements $\bar{\xi}_1, \cdots, \bar{\xi}_r$ have the property that they are algebraically independent[6] and that every element in $\mathfrak{o}$ is integrally dependent on $\bar{\xi}_1, \cdots, \bar{\xi}_r$. Hence our definition of a simple point P can now be completed by the following remark: the r elements $\eta_1, \cdots, \eta_r$ enjoying the property that $\mathfrak{p}_0[= (\xi_1, \cdots, \xi_n)]$

---

[5] See [2], p. 41.

[6] The algebraic independence is already implied by the non-vanishing of the determinant $| c_{ij} |$, as a consequence of the fact that $\mathfrak{p}_0$ is then an isolated component of the ideal $(\bar{\xi}_1, \cdots, \bar{\xi}_r)$. Namely, if the $\bar{\xi}$'s are algebraically dependent, it is permissible to assume that $\bar{\xi}_r$ is integrally dependent on $\bar{\xi}_1, \cdots, \bar{\xi}_{r-1}$. In view of the algebraic closure of the ground field, this implies that if $\mathfrak{p}$ is any minimal prime ideal of the ideal $(\bar{\xi}_1, \cdots, \bar{\xi}_{r-1})$ which is a multiple of $\mathfrak{p}_0$, then $\bar{\xi}_r$, and hence also the ideal $(\bar{\xi}_1, \cdots, \bar{\xi}_r)$, is divisible by $\mathfrak{p}$.

(the point P being at the origin) is an isolated component of the ideal $(\eta_1, \cdots, \eta_r)$, can be chosen in such a manner that, in addition, every element in $\mathfrak{o}$ be integrally dependent on $\eta_1, \cdots, \eta_r$; in particular, r suitable linear forms in $\xi_1, \cdots, \xi_n$ will meet both these requirements. We make the necessary transformation of the coördinates ξ_i and we assume from now on that these linear forms are $\xi_1, \cdots, \xi_r$ respectively.

3. Let again $\eta_1, \cdots, \eta_r$ be r elements in $\mathfrak{o}$ with the property that $\mathfrak{p}_0$ is an isolated component of the ideal $(\eta_1, \cdots, \eta_r)$. As has been pointed out above (footnote 6), the η's are algebraically independent. We introduce the ring $\mathsf{K}\{\eta\} = \mathsf{K}\{\eta_1, \cdots, \eta_r\}$ of all formal power seriès in $\eta_1, \cdots, \eta_r$ with coefficients in K. Any element in $\mathsf{K}\{\eta\}$ can be written in the form

$$\psi_0 + \psi_1 + \cdots + \psi_m + \cdots, \quad \text{or} \quad \phi_m + R_{m+1},$$

where ψ_i is a form of degree i in $\eta_1, \cdots, \eta_r$ and where $\phi_m = \psi_0 + \psi_1 + \cdots + \psi_m$.

THEOREM 2. *There exists an isomorphic mapping of the ring* $\mathfrak{o}$ *upon a subring of* $\mathsf{K}\{\eta\}$,[7] *with the following property: if* ω *is any element in* $\mathfrak{o}$ *and if* $\psi_0 + \psi_1 + \cdots$ *is the corresponding element in* $\mathsf{K}\{\eta\}$, *then, for all* m,

$$\omega \equiv \phi_m(\mathfrak{p}_0^{m+1}), \qquad \phi_m = \psi_0 + \psi_1 + \cdots + \psi_m. \tag{6}$$

Proof. We shall first prove the theorem in the case when $\eta_1, \cdots, \eta_r$ coincide with $\xi_1, \cdots, \xi_r$ respectively. Its validity in the general case will then be an immediate consequence.

We first show that given any element ω in $\mathfrak{o}$ there exists a polynomial $\phi_m(\xi_1, \cdots, \xi_r)$ of degree $\leqq m$, such that the congruence (6) holds true. The assertion is trivial for $m = 0$, and has been proved for $m = 1$ (see congruence (5)). We assume that the assertion is true for $m = i$ and we prove it for $m = i + 1$. Let then $\omega \equiv \phi_i(\mathfrak{p}_0^{i+1})$, where $\phi_i = \phi_i(\xi_1, \cdots, \xi_r)$ is a polynomial of degree $\leqq i$. We can write:

$$\omega = \phi_i + f_{i+1}(\xi_1, \cdots, \xi_n) + \cdots + f_m(\xi_1, \cdots, \xi_n),$$

where f_j is a form of degree j. Let

$$\xi_{r+i} \equiv \sum_{j=1}^{r} c_{ij}\xi_j(\mathfrak{p}_0^2), \qquad (i = 1, 2, \cdots, n - r).$$

---

[7] In a more precise language: there exists an isomorphic mapping of $\mathfrak{o}$ upon a subring of $\mathsf{K}\{u_1, \cdots, u_r\}$ (the u's being parameters) in which $\eta_i \to u_i$, etc. What we did is to identify the η's with the u's.

Then it is clear that

$$f_{i+1}(\xi_1,\cdots,\xi_n)\equiv f_{i+1}(\xi_1,\cdots,\xi_r,\sum_{j=1}^{r}c_{1j}\xi_j,\cdots,\sum_{j=1}^{r}c_{n-r,j}\xi_j)\,(\mathfrak{p}_0^{i+2}),$$

and that moreover $f_j(\xi_1,\cdots,\xi_n)\equiv 0(\mathfrak{p}_0^{i+2})$, if $j>i+1$. Hence, if we put

$$(7)\qquad \begin{cases}\psi_{i+1}=f_{i+1}(\xi_1,\cdots,\xi_r,\sum_{j=1}^{r}c_{1j}\xi_j,\cdots,\sum_{j=1}^{r}c_{n-r,j}\xi_j)\\ \phi_{i+1}=\phi_i+\psi_{i+1},\end{cases}$$

then

$$\omega\equiv\phi_{i+1}(\mathfrak{p}_0^{i+2}),$$

which proves our assertion.

We next show that, given the element ω, the polynomial ϕ_m in (6) is uniquely determined. We shall prove this by induction with respect to m, since for $m=0$ the assertion is trivial. Let us assume that there exist two polynomials, $\phi_m(\xi_1,\cdots,\xi_r)$ and $\phi'_m(\xi_1,\cdots,\xi_r)$, both of degree $\leqq m$, such that $\omega\equiv\phi_m(\mathfrak{p}_0^{m+1})$, $\omega\equiv\phi'_m(\mathfrak{p}_0^{m+1})$. Let $\phi_m=\phi_{m-1}+\psi_m$, $\phi'_m=\phi'_{m-1}+\psi'_m$, where ϕ_{m-1} and ϕ'_{m-1} are of degree $\leqq m-1$ and ψ_m and ψ'_m are forms of degree m. Since ψ_m and ψ'_m are in $\mathfrak{p}_0^m$, it follows that $\omega\equiv\phi_{m-1}(\mathfrak{p}_0^m)$ and $\omega\equiv\phi'_{m-1}(\mathfrak{p}_0^m)$. By our induction we conclude that $\phi_{m-1}=\phi'_{m-1}$. Hence $\phi_m-\phi'_m=\psi_m-\psi'_m=g_m$, and $g_m\equiv 0(\mathfrak{p}_0^{m+1})$, where g_m is a form in $\xi_1,\cdots,\xi_r$, not identically zero, of degree m.

Let us denote the ideal $(\xi_1,\cdots,\xi_r)$ by $\mathfrak{A}$. By hypothesis, $\mathfrak{p}_0$ is an isolated component of $\mathfrak{A}$. Moreover, since every element in $\mathfrak{o}$ is integrally dependent on $\xi_1,\cdots,\xi_r$, $\mathfrak{A}$ is unmixed, all its components being necessarily zero-dimensional. Let $\mathfrak{A}=[\mathfrak{p}_0,\mathfrak{q}_1,\cdots,\mathfrak{q}_s]$. Since all the components are zero-dimensional, no two of them have common divisors. Hence their intersection coincides with their product, and we can write $\mathfrak{A}=\mathfrak{p}_0\mathfrak{q}_1\cdots\mathfrak{q}_s$. Now $g_m=0(\mathfrak{A}^m)$, i. e. $g_m\subset\mathfrak{p}_0^m\mathfrak{q}_1^m\cdots\mathfrak{q}_s^m=[\mathfrak{p}_0^m,\mathfrak{q}_1^m,\cdots,\mathfrak{q}_s^m]$. On the other hand $g_m\equiv 0(\mathfrak{p}_0^{m+1})$, whence $g_m\subset[\mathfrak{p}_0^{m+1},\mathfrak{q}_1^m,\cdots,\mathfrak{q}_s^m]=\mathfrak{p}_0^{m+1}\mathfrak{q}_1^m\cdots\mathfrak{q}_s^m$. Hence

$$g_m\equiv 0(\mathfrak{p}_0\cdot\mathfrak{A}^m).$$

This last congruence shows that g_m can be expressed as a form $f(\xi_1,\cdots,\xi_r)$ in $\xi_1,\cdots,\xi_r$, of degree m, *with coefficients in* $\mathfrak{p}_0$. Consider the norm $N(f)$ of f over $\mathsf{K}(\xi_1,\cdots,\xi_r)$. Clearly $N(f)=g_m^\nu$, where $\nu=[\Sigma:\mathsf{K}(\xi_1,\cdots,\xi_r)]$, the relative degree of Σ over $\mathsf{K}(\xi_1,\cdots,\xi_r)$. On the other hand $N(f)$ is a form of degree νm in $\xi_1,\cdots,\xi_r$, and its coefficients are polynomials in $\xi_1,\cdots,\xi_r$ (since the coefficients of f are integrally dependent on $\xi_1,\cdots,\xi_r$) which vanish for $\xi_1=\cdots=\xi_r=0$ (since the coefficients of f are in $\mathfrak{p}_0$).[8]

---

[8] All this can be seen as follows. Let $u_1,u_2,\cdots,u_M$ be the various power products $\xi_1^{i_1}\cdots\xi_r^{i_r}$, $i_1+i_2+\cdots+i_r=m$. Then f is of the form $f=u_1\alpha_1+u_2\alpha_2+\cdots$

We have therefore the following result: the form $g_m{}^\nu$ in $\xi_1, \cdots, \xi_r$, of degree $m\nu$, is equal to a polynomial in $\xi_1, \cdots, \xi_r$, having no terms of degree $< m\nu + 1$. This is in contradiction with the algebraic independence of $\xi_1, \cdots, \xi_r$. The uniqueness of the polynomial ϕ_m in (6) is thus established.

The formula (7) shows that if $\omega \equiv \phi_m(\mathfrak{p}_0{}^{m+1})$ and $\omega \equiv \phi_{m+1}(\mathfrak{p}_0{}^{m+2})$, then $\phi_{m+1} - \phi_m$ is a form ψ_{m+1} in $\xi_1, \cdots, \xi_r$, of degree $m+1$. *We let correspond to* ω *the power series* $\psi = \psi_0 + \psi_1 + \cdots$:

$$\omega \to \psi = \sum_{i=0}^{\infty} \psi_i(\xi_1, \cdots, \xi_r). \tag{8}$$

It is now a simple matter to show that (8) defines an isomorphic mapping of $\mathfrak{o}$ upon a subring H of $K\{\xi_1, \cdots, \xi_r\}$.

(a) First, let ω and ω' be distinct elements of $\mathfrak{o}$, and let

$$\omega \to \psi = \sum_{i=0}^{\infty} \psi_i, \qquad \omega' \to \psi' = \sum_{i=0}^{\infty} \psi'_i.$$

Assume that $\psi_i = \psi'_i$, for all i. Since

$$\omega - \sum_{i=0}^{m} \psi_i \equiv 0(\mathfrak{p}_0{}^{m+1}) \quad \text{and} \quad \omega' - \sum_{i=0}^{m} \psi'_i \equiv 0(\mathfrak{p}_0{}^{m+1}),$$

it would follow that $\omega - \omega' \equiv 0(\mathfrak{p}_0{}^{m+1})$, for all m. This is impossible if $\omega \neq \omega'$, since zero is the only element which is common to all the powers of $\mathfrak{p}_0$. Hence, *to two distinct elements of* $\mathfrak{o}$ *there correspond two distinct power series.*

(b) We have $\omega + \omega' - \sum_{i=0}^{m} (\psi_i + \psi'_i) \equiv 0(\mathfrak{p}_0{}^{m+1})$, hence

$$\omega + \omega' \to \sum_{i=0}^{\infty} (\psi_i + \psi'_i) = \psi + \psi'.$$

$+ u_M a_M$, $a_j \equiv \mathfrak{o}(\mathfrak{p}_0)$, $j = 1, 2, \cdots, M$. Let $f^{(1)}, f^{(2)}, \cdots, f^{(\nu-1)}$ be the conjugates of f over $K(\xi_1, \cdots, \xi_r)$; $f^{(i)} = u_1 a_1{}^{(i)} + u_2 a_2{}^{(i)} + \cdots + u_M a_M{}^{(i)}$, where $a_j{}^{(1)}, \cdots, a_j{}^{(\nu-1)}$ are the conjugates of a_j. The conjugates of any a_j are integrally dependent on $\xi_1, \cdots, \xi_r$. Hence $N(f)$ $(= ff^{(1)} \cdots f^{(\nu-1)})$ is a form of degree νm in $\xi_1, \cdots, \xi_r$ with coefficients which are polynomials in $\xi_1, \cdots, \xi_r$. We consider the least Galois extension of $K(\xi_1, \cdots, \xi_r)$ which contains the field Σ. We also consider the smallest ring $\mathfrak{o}^*$ containing $\mathfrak{o}$ and its conjugate rings, i.e. the ring $\mathfrak{o}^* = K[\xi_1, \cdots, \xi_n;\ \xi_1{}^{(1)}, \cdots, \xi_n{}^{(1)};\ \cdots;\ \xi_1{}^{(\nu-1)}, \cdots, \xi_n{}^{(\nu-1)}]$. Let $\mathfrak{p}_0{}^*$ be any prime ideal divisor of $\mathfrak{p}_0$ in $\mathfrak{o}^*$. The coefficients of $N(f)$, considered as a form of degree νm in $\xi_1, \cdots, \xi_r$, clearly belong to the ideal $\mathfrak{o}^*(a_1, \cdots, a_M)$. Since $a_j \equiv 0(\mathfrak{p}_0) \equiv 0(\mathfrak{p}_0{}^*)$, all these coefficients belong to $\mathfrak{p}_0{}^*$. On the other hand, these coefficients are in $K[\xi_1, \cdots, \xi_r]$, and it is clear that the intersection $K[\xi_1, \cdots, \xi_r] \cap \mathfrak{p}^*_0$ is the prime 0-dimensional ideal $(\xi_1, \cdots, \xi_r)$. Hence the coefficients of $N(f)$ are polynomials in $\xi_1, \cdots, \xi_r$, without constant term.

(c) Let $\psi\psi' = g = g_0 + g_1 + \cdots$. Clearly, $\sum_{i=0}^{m} g_i$ differs from $\sum_{i=0}^{m} \psi_i \cdot \sum_{i=0}^{m} \psi'_i$ by terms of degree $> m$, whence

$$\sum_{i=0}^{m} \psi_i \cdot \sum_{i=0}^{m} \psi'_i - \sum_{i=0}^{m} g_i \equiv 0(\mathfrak{p}_0^{m+1}). \tag{9}$$

On the other hand, we have

$$\omega\omega' - \sum_{i=0}^{m} \psi_i \cdot \sum_{i=0}^{m} \psi'_i = (\omega - \sum_{i=0}^{m} \psi_i)\omega' + (\omega' - \sum_{i=0}^{m} \psi'_i) \sum_{i=0}^{m} \psi_i,$$

whence

$$\omega\omega' - \sum_{i=0}^{m} \psi_i \cdot \sum_{i=0}^{m} \psi'_i \equiv 0(\mathfrak{p}_0^{m+1}). \tag{9'}$$

Adding (9) and (9′), we obtain:

$$\omega\omega' - \sum_{i=0}^{m} g_i \equiv 0(\mathfrak{p}_0^{m+1}),$$

and this shows that to $\omega\omega'$ there corresponds the power series $\psi\psi'$. This completes the proof of Theorem 2 for the case when $\eta_1, \cdots, \eta_r$ coincide with $\xi_1, \cdots, \xi_r$.

The general case can now be settled in a few words. Let us consider the power series which correspond to $\eta_1, \cdots, \eta_r$ in our mapping of $\mathfrak{o}$ upon the subring H of $\mathsf{K}\{\xi_1, \cdots, \xi_r\}$. The constant term in each of these power series is zero, since $\eta_i \equiv 0(\mathfrak{p}_0)$. Let

(10) $\eta_i \rightarrow \eta^*_i = c_{i1}\xi_1 + \cdots + c_{ir}\xi_r +$ terms of higher degree, $\eta^*_i \subset H$, $(i = 1, 2, \cdots, r)$.

Since $\mathfrak{p}_0$ is an isolated component of the ideal $(\eta_1, \cdots, \eta_r)$, it follows by Theorem 1, that the determinant $| c_{ij} |$ is different from zero. It is therefore possible to "solve" the relations (10) for $\xi_1, \cdots, \xi_r$ and to express each ξ_i as a power series in $\eta^*_1, \cdots, \eta^*_r$. In other terms: the ring $\mathsf{K}\{\xi_1, \cdots, \xi_r\}$ can be mapped isomorphically upon the ring $\mathsf{K}\{\eta_1, \cdots, \eta_r\}$. In this mapping the subring H of $\mathsf{K}\{\xi_1, \cdots, \xi_r\}$ is mapped upon a subring H' of $\mathsf{K}\{\eta_1, \cdots, \eta_r\}$ and $\mathfrak{o} \cong H'$. Let ω be any element in $\mathfrak{o}$ and let

$$\omega \rightarrow \sum_{i=0}^{\infty} g_i = g, \qquad g \subset H',$$

where g_i is a form in $\eta_1, \cdots, \eta_r$ of degree i. The two power series in

$\mathsf{K}\{\xi_1,\cdots,\xi_r\}$ which correspond to ω and to $\sum_{i=0}^{m} g_i$ respectively, differ by terms of degree $> m$. Hence $\omega - \sum_{i=0}^{m} g_i \equiv 0(\mathfrak{p}_0^{m+1})$, q. e. d.

In the sequel we shall feel free to identify any element ω with its corresponding power series $\sum_{i=0}^{\infty} g_i$ in $\eta_1,\cdots,\eta_r$, and we shall write $\omega = \sum_{i=0}^{\infty} g_i$. We shall say that $\sum_{i=0}^{\infty} g_i$ is *the expansion of* ω *in a power series of* $\eta_1,\cdots,\eta_r$. We shall also refer to the isomorphic mapping of $\mathfrak{o}$ upon the subring H of $\mathsf{K}\{\eta_1,\cdots,\eta_r\}$ as a *uniformization at* P. The elements $\eta_1,\cdots,\eta_r$ shall be called *uniformizing parameters.*

4. Geometrically speaking, Theorem 2, in the special case $\eta_i = \xi_i$, says that the variety V_r possesses at the point $P(0,\cdots,0)$ a linear analytical branch, whose tangent S_r at P has only the point P in common with the S_{n-r} given by the equation $\xi_1 = \cdots = \xi_r = 0$. We have not yet shown that there are no other branches at P, and that consequently P is indeed a simple point in the ordinary geometric sense.[9]

[9] Actually the existence of a linear analytical branch at P is implied already by a part of Theorem 2. Namely, if we only knew that there exists an isomorphic mapping of $\mathfrak{o}$ upon a subring of $\mathsf{K}\{\xi_1,\cdots,\xi_r\}$ such that the congruence (6) holds true *for* $m = 0$, we could already associate with this mapping a linear branch at $P(0,\cdots,0)$. It will follow from Theorem 4 that it is the validity of the congruence (6) *for all* m that implies that there are no other branches through P. Thus *the existence of an isomorphic mapping of* $\mathfrak{o}$ *upon a subring of* $\mathsf{K}\{\xi_1,\cdots,\xi_r\}$ *together with the validity of the congruence* (6) *for all* m *characterizes a simple point of* V_2. In this connection we wish to call attention to the following question: *assuming* $\mathfrak{o}$ *to be integrally closed* (i. e. assuming that V_2 is *normal* in the affine space (section 14)), *is it true that the neighborhood of every point of* V_2 *is an analytically irreducible variety?* For algebraic functions of one variable the answer is well-known to be affirmative. It seems to us that for functions of several variables the question presents serious difficulties. An equivalent ideal theoretic formulation of the question is the following: Let ω be a primitive element of $\mathfrak{o}$ over $\mathsf{K}(\xi_1,\cdots,\xi_r)$, $\mathfrak{o}$ being assumed to be the integral closure of $\mathsf{K}[\xi_1,\cdots,\xi_r]$ in Σ. Let $F(\xi_1,\cdots,\xi_r,\omega) = 0$ be the defining (irreducible) equation for ω. Over the field of meromorphic functions of $\xi_1,\cdots,\xi_r$ the polynomial F factors: $F_1F_2\cdots F_h$, where F_h is a polynomial in ω with coefficients in $\mathsf{K}\{\xi_1,\cdots,\xi_r\}$ and leading coefficient 1. It is not difficult to associate with each factor F_i a definite prime 0-dimensional ideal in $\mathfrak{o}$ which is a divisor of $\mathfrak{o}\cdot(\xi_1,\cdots,\xi_r)$ (see next section). The question is: do there correspond to distinct factors distinct ideals? The methods used in this connection for algebraic functions of one variable and for fields of algebraic numbers (see [4]) break down in the case of algebraic functions of more than one variable. If the answer to this question was affirmative, then the following could be proved: *if* $\mathfrak{o}$ *is integrally closed and if there exists an isomorphic mapping of* $\mathfrak{o}$ *upon a subring of* $\mathsf{K}\{\xi_1,\cdots,\xi_r\}$ *such that* (6) *holds true for* $m = 0$, *then* (6) *holds true for*

We will show, however, that there exists a non-singular projection V'_r of V_r into an $S_{r+1}(y_1, \cdots, y_{r+1})$, such that if $f(y_1, \cdots, y_{r+1}) = 0$ is the equation of V'_r, then the point P is projected into a point P' of V'_r at which not all the derivatives f_{y_i} vanish and which is therefore a simple point of V'_r, in the ordinary sense of this term.

In addition to this we shall also wish to show that the existence of a uniformization at P with the property described in Theorem 2 (i. e. congruence (6), for all m) implies that P is a simple point of V_r (in the sense of our definition). Accordingly, we consider a point P of V_r, which we shall assume to be the origin (whence the corresponding prime 0-dimensional ideal in $\mathfrak{o}$ is $\mathfrak{p}_0 = (\xi_1, \cdots, \xi_n)$) and we make the following assumptions: *there exist in* $\mathfrak{o}$ *r uniformizing parameters* $\eta_1, \cdots, \eta_r$ for the point P, i. e. there exists an isomorphic mapping of $\mathfrak{o}$ upon a subring of $\mathsf{K}\{\eta_1, \cdots, \eta_r\}$ in which every element of $\mathfrak{p}_0$ is mapped upon a power series in $\eta_1, \cdots, \eta_r$ which has no constant term (congruence (6), for $m = 0$!). We also assume that *this uniformization has the property that if to an element* ω *in* $\mathfrak{o}$ *there corresponds the power series* $\psi_0 + \psi_1 + \cdots$, *then* $\omega \equiv \psi_0 + \cdots + \psi_m (\mathfrak{p}_0^{m+1})$, *for all* m (congruence (6), for all m). When these assumptions are satisfied, we shall say briefly that *there exists a uniformization of the whole neighborhood of* P.

THEOREM 3. *If* $\eta_1, \cdots, \eta_r$ *are uniformizing parameters for a uniformization of the whole neighborhood of the point* $P(0, \cdots, 0)$, *then* $\mathfrak{p}_0$ *is an isolated component of the ideal* $\mathfrak{o} \cdot (\eta_1, \cdots, \eta_r)$ *and hence* P *is a simple point of* V_r.

Proof. By hypothesis, any element ω in $\mathfrak{o}$ satisfies a congruence of the form $\omega \equiv c_0 + c_1\eta_1 + \cdots + c_r\eta_r (\mathfrak{p}_0^2)$, and $c_0 = 0$ if $\omega \equiv 0 (\mathfrak{p}_0)$. It follows immediately that $\mathfrak{p}_0 = (\eta_1, \cdots, \eta_r, \mathfrak{p}_0^2)$, and this implies (see proof of Theorem 1) that $\mathfrak{p}_0$ is an isolated component of the ideal $\mathfrak{o} \cdot (\eta_1, \cdots, \eta_r)$, q. e. d.

We observe that we have used in the proof only the congruence $\omega \equiv \psi_0 + \psi_1 (\mathfrak{p}_0^2)$. We can therefore state the preceding theorem in the following stronger form:

THEOREM 3. 1. *If there exist r elements* $\eta_1, \cdots, \eta_r$ *in* $\mathfrak{o}$ *such that every element* ω *in* $\mathfrak{o}$ *satisfies a congruence of the form* $\omega \equiv c_0 + c_1\eta_1 + \cdots + c_r\eta_r (\mathfrak{p}_0^2)$ (*congruence* (6), *for* $m = 0, 1$ *only*!), *then* $\mathfrak{p}_0$ *is an isolated component of the ideal* $\mathfrak{o} \cdot (\eta_1, \cdots, \eta_r)$ *and* P *is a simple point of* V_r.

---

all m, and hence P is a simple point. Geometrically speaking, this statement signifies just this: the analytical irreducibility of the neighborhood of P together with the existence of a linear branch through P characterize P as a simple point.

The assumption in Theorem 3.1 implies that the elements $\eta_1, \cdots, \eta_r$ form a modular basis of the ring $\mathfrak{p}_0/\mathfrak{p}_0^2$ considered as a K-module. Moreover, since $\mathfrak{p}_0$ is, according to this theorem, an isolated component of $\mathfrak{o} \cdot (\eta_1, \cdots, \eta_r)$, we know, from section 2, that $\eta_1, \cdots, \eta_r$ are linearly independent modulo $\mathfrak{p}_0^2$. *Hence $\mathfrak{p}_0/\mathfrak{p}_0^2$ is a K-module of rank r.* Conversely, if $\mathfrak{p}_0/\mathfrak{p}_0^2$ is of rank $\leqq r$, then there exist r elements in $\mathfrak{o}$, say $\eta_1, \cdots, \eta_r$, such that every other element ω in $\mathfrak{p}_0$ satisfies a congruence of the form $\omega \equiv c_1\eta_1 + \cdots + c_r\eta_r(\mathfrak{p}_0^2)$. Hence $\mathfrak{p}_0 = (\eta_1, \cdots, \eta_r, \mathfrak{p}_0^2)$, and from this it follows, as in section 2, that $\mathfrak{p}_0$ is an isolated component of the ideal $\mathfrak{o} \cdot (\eta_1, \cdots, \eta_r)$ and that $\eta_1, \cdots, \eta_r$ are linearly independent mod $\mathfrak{p}_0^2$. Hence P is a simple point and necessarily $\mathfrak{p}_0/\mathfrak{p}_0^2$ is of rank exactly r. We therefore have the following

THEOREM 3.2. *If P is a point on V_r and $\mathfrak{p}_0$ is the corresponding prime 0-dimensional ideal in $\mathfrak{o}$, a necessary and sufficient condition that P be a simple point is that the K-module $\mathfrak{p}_0/\mathfrak{p}_0^2$ be of rank r.*

5. Let $P(0, \cdots, 0)$ be a simple point of V_r. We assume, as in the preceding section, that a linear homogeneous transformation of the coördinates ξ_i has already been performed, so as to make $\xi_1, \cdots, \xi_r$ uniformizing parameters for a uniformization of the whole neighborhood of the point P, and, moreover, that every element in $\mathfrak{o}$ is integrally dependent on $K[\xi_1, \cdots, \xi_r]$.

The field $\Sigma = K(\xi_1, \cdots, \xi_n)$ is an algebraic extension of $K(\xi_1, \cdots, \xi_r)$. Let m be the relative degree of Σ over $K(\xi_1, \cdots, \xi_r)$. If ω is an element in $\mathfrak{o}$, then

$$N(z - \omega) = F(z; \xi_1, \cdots, \xi_r) = z^m + A_1 z^{m-1} + \cdots + A_m,$$

where $A_1, \cdots, A_m$ are polynomials in $\xi_1, \cdots, \xi_r$. Reducing the equation $F(\omega; \xi_1, \cdots, \xi_r) = 0$ modulo $\mathfrak{p}_0$, we find $F(c, 0, \cdots, 0) = 0$, where $\omega \equiv c(\mathfrak{p}_0)$. Hence c is a root of $F(z, 0, \cdots, 0)$.

Since every element in $\mathfrak{o}$ is integrally dependent on $K[\xi_1, \cdots, \xi_r]$, the ideal $\mathfrak{o} \cdot (\xi_1, \cdots, \xi_r)$ is unmixed and zero-dimensional, and $\mathfrak{p}_0$ is one of its components. Let $[\mathfrak{p}_0, \mathfrak{q}_1, \cdots, \mathfrak{q}_s]$ be the decomposition of the ideal $\mathfrak{o} \cdot (\xi_1, \cdots, \xi_r)$ into primary components and let $\mathfrak{p}_0, \mathfrak{p}_1, \cdots, \mathfrak{p}_s$ be the associated prime zero-dimensional ideals.

THEOREM 4. *Under the assumption that $\xi_1, \cdots, \xi_r$ are uniformizing parameters for a uniformization of the whole neighborhood of the simple point $P(0, \cdots, 0)$ on V_r and that every element in $\mathfrak{o}$ is integrally dependent on $\xi_1, \cdots, \xi_r$, there exist elements ω in $\mathfrak{o}$ such that $F'_\omega(\omega; \xi_1, \cdots, \xi_r) \not\equiv 0(\mathfrak{p}_0)$, where $F(z; \xi_1, \cdots, \xi_r)$ is the norm of $z - \omega$ over $K(\xi_1, \cdots, \xi_r)$. Such elements ω are characterized by the following condition: if ω has the*

value c at P (i. e. *if* $\omega \equiv c(\mathfrak{p}_0)$) *and if* $\mathfrak{p}_1, \cdots, \mathfrak{p}_s$ *are the prime* 0-*dimensional ideals, other than* $\mathfrak{p}_0$, *which divide the ideal* $\mathfrak{o} \cdot (\xi_1, \cdots, \xi_r)$, *then* $\omega \not\equiv c(\mathfrak{p}_i)$, $i = 1, 2, \cdots, s$.

We give here two proofs of this theorem. The first is based upon the *factorization of* $F(z; \xi_1, \cdots, \xi_r)$ in the ring $K\{\xi_1, \cdots, \xi_r\}[z]$ and upon higher congruences to which this factorization leads.[10]

The second proof makes use of the least Galois extension field which contains our field Σ and its conjugates over $K(\xi_1, \cdots, \xi_r)$.

First proof. Elements ω in $\mathfrak{o}$ satisfying the conditions $\omega \equiv c(\mathfrak{p}_0)$, $\omega \not\equiv c(\mathfrak{p}_i)$, $i = 1, 2, \cdots, s$, certainly exist. Namely, since no two of the ideals $\mathfrak{p}_0, \mathfrak{p}_1, \cdots, \mathfrak{p}_s$ have common divisors, it is well known that given any $s + 1$ elements $\alpha_0, \alpha_1, \cdots, \alpha_s$ in $\mathfrak{o}$, there exist elements ω in $\mathfrak{o}$ satisfying the congruences $\omega \equiv \alpha_i(\mathfrak{p}_i)$, $i = 0, 1, \cdots, s$. To satisfy the above conditions, we have only to take for $\alpha_0, \cdots, \alpha_s$ any $s + 1$ constants in K, say $c_0, c_1, \cdots, c_s$, such that $c_0 \neq c_i$, $i = 1, 2, \cdots, s$.

It is not difficult to prove that *there also exist primitive elements* ω *over* $K(\xi_1, \cdots, \xi_r)$ satisfying the above conditions (as a matter of fact, we shall prove later on that *these conditions imply that* ω *is a primitive element*). To show this we consider an element ζ in $\mathfrak{o}$ satisfying the congruences $\zeta \equiv d_i(\mathfrak{p}_i)$, where $d_0, d_1, \cdots, d_s$ are *distinct* constants. Let ζ' be some primitive element in $\mathfrak{o}$ and let $\zeta' \equiv d'_i(\mathfrak{p}_i)$. If t is a parameter, then the discriminant of $N(z - \zeta' - t\zeta)$ is a polynomial $D(t; \xi_1, \cdots, \xi_r)$ which does not vanish identically in t, since $D(0, \xi_1, \cdots, \xi_r)$ is the discriminant of $N(z - \zeta')$ and does not vanish. Let $A_{ij}(t) = (d'_i - d'_j) + t(d_i - d_j)$, $i, j = 0, 1, 2, \cdots, s$, $i \neq j$. Since $d_i \neq d_j$, if $i \neq j$, none of these polynomials vanishes identically. Let t_0 be a value of t (in K) such that $D(t_0, \xi_1, \cdots, \xi_r) \neq 0$, $A_{ij}(t_0) \neq 0$. Then the element $\omega = \zeta' + t_0\zeta$ is obviously a primitive element. Moreover, we have $\omega \equiv d'_i + t_0 d_i(\mathfrak{p}_i)$ and $d'_i + t_0 d_i \neq d'_j + t_0 d'_j$, if $i \neq j$, since $A_{ij}(t_0) \neq 0$.

Let then ω be a primitive element in $\mathfrak{o}$ over $K(\xi_1, \cdots, \xi_r)$, $\omega \equiv c_i(\mathfrak{p}_i)$, $c_i \neq c_0$, $i = 1, 2, \cdots, s$. Let

$$N(z - \omega) = F(z; \xi_1, \cdots, \xi_r) = F(z; \xi).$$

[10] This proof follows in part a well-known pattern of the theory of algebraic functions of one variable and of algebraic number fields. See [4]. The generalization is made possible by the existence of the uniformization. In the case of algebraic functions of one variable and of algebraic number fields the proof is based upon the fact that the ring of polynomials in one variable and the ring of rational integers are principal ideal rings.

Here F is a polynomial in z and $\xi_1, \cdots, \xi_r$, of degree m ($= [\Sigma : \mathsf{K}(\xi_1, \cdots, \xi_r)]$) in z and with leading coefficient 1.

Let

$$F = F_1(z;\xi) \cdot F_2(z;\xi) \cdots F_h(z;\xi) \tag{11}$$

be the decomposition of F into prime factors in the ring $\mathsf{K}\{\xi_1, \cdots, \xi_r\}[z]$ of polynomials in z with coefficients in the ring of formal power series in $\xi_1, \cdots, \xi_r$. Since the discriminant of F is different from zero, the prime factors F_i are distinct. Moreover, each factor F_i, of degree m_i in z, may be assumed to have leading coefficient 1. We have $m = m_1 + \cdots + m_h$.

Since $F_j(z;\xi)$ is a prime element in $\mathsf{K}\{\xi\}[z]$, it follows by Weierstrass' preparation theorem [11] that the polynomial $F_j(z; 0, 0, \cdots, 0)$ in $\mathsf{K}[z]$ cannot have two distinct roots. Hence

$$F_j(z, 0, \cdots, 0) = (z - d_j)^{m_j}, \qquad d_j \subset \mathsf{K}. \tag{12}$$

The h constants $d_1, d_2, \cdots, d_h$ (not necessarily distinct) are the roots of $F(z, 0, \cdots, 0)$:

$$F(z; 0, \cdots, 0) = \prod_{j=1}^{h} (z - d_j)^{m_j}. \tag{12'}$$

Every one of the constants c_i, $i = 0, 1, \cdots, s$, is equal to some d_j. In fact, reducing the equation $F(\omega; \xi_1, \cdots, \xi_r)$ modulo $\mathfrak{p}_i$, we find $F(c_i; 0, \cdots, 0) = 0$, i. e. c_i is a root of $F(z; 0, \cdots, 0)$. *We shall have occasion to prove later* (footnote 12) *that, conversely, each of the constants d_j is equal to some c_i.* At any rate, if the constants $c_0, c_1, \cdots, c_s$ were distinct, then necessarily $h \geqq s + 1$.

Let $\omega = \phi = \phi_0 + \phi_1 + \cdots$ be the expansion of ω into a power series in $\xi_1, \cdots, \xi_r$. Here $\phi_0 = c_0$. Since the mapping $\omega \to \phi$ of $\mathfrak{o}$ upon the subring H of $\mathsf{K}\{\xi_1, \cdots, \xi_r\}$ is an isomorphism, the equation $F(\omega; \xi) = 0$ implies that ϕ is a root of $F(z;\xi)$. Hence $z - \phi$ is a factor of F. Let, say

$$F_1(z;\xi) = z - \phi,$$

whence $m_1 = 1$ and

$$F(z, 0, \cdots, 0) = (z - c_0) \prod_{j=2}^{h} (z - d_j)^{m_j}, \quad m_2 + \cdots + m_h = m - 1. \tag{13}$$

We wish to show that the constants $d_2, \cdots, d_h$ are all different from c_0, i. e. that c_0 is a simple root of $F(z, 0, \cdots, 0)$. This will imply that

[11] See [5], p. 261-262.

$$F'_\omega(\omega; \xi_1, \cdots, \xi_r) \not\equiv 0(\mathfrak{p}_0),$$

as was asserted in our theorem. For the proof, let us denote by $F_{j\rho}(z;\xi)$ the polynomial in z and in $\xi_1, \cdots, \xi_r$ obtained from $F_j(z;\xi)$ by omitting all terms of degree $> \rho$ in $\xi_1, \cdots, \xi_r$. It is clear that $F(z;\xi) - \prod_{j=1}^{h} F_{j\rho}(z;\xi)$ is a polynomial which does not contain terms of degree $< \rho + 1$ in $\xi_1, \cdots, \xi_r$. It follows that

(14) $$\prod_{j=1}^{h} F_{j\rho}(\omega;\xi) \equiv 0(\xi_1, \cdots, \xi_r)^{\rho+1},$$

or also, since $\mathfrak{o} \cdot (\xi_1, \cdots, \xi_r) = \mathfrak{p}_0\mathfrak{q}_1 \cdots \mathfrak{q}_s$,

(14′) $$\prod_{j=1}^{h} F_{j\rho}(\omega;\xi) \equiv 0(\mathfrak{p}_0^{\rho+1}\mathfrak{q}_1^{\rho+1} \cdots \mathfrak{q}_s^{\rho+1}).$$

Now suppose that one of the constants $d_2, \cdots, d_h$ coincides with $c_0(= d_1)$, and let, for instance $d_h = c_0$. We show that this hypothesis leads to a contradiction. It is clear that $F_{j\rho}(z; 0, \cdots, 0) = (z - d_j)^{m_j}$. Hence $F_{j\rho}(\omega;\xi) \equiv (\omega - d_j)^{m_j}(\mathfrak{p}_i)$, $i = 0, 1, \cdots, s$. In particular, for $j = h$, we have

$$F_{h\rho}(\omega;\xi) \equiv (c_i - d_h)^{m_h}(\mathfrak{p}_i), \qquad (i = 0, 1, \cdots, s).$$

Since $d_h = c_0$ and $c_0 \neq c_i$, $i = 1, 2, \cdots, s$, we conclude that

(15) $$F_{h\rho}(\omega;\xi) \not\equiv 0(\mathfrak{p}_i), \qquad (i = 1, 2, \cdots, s).$$

From this we derive, in view of (14′), that

$$\prod_{j=1}^{h-1} F_{j\rho}(\omega;\xi) \equiv 0(\mathfrak{q}_1^{\rho+1} \cdots \mathfrak{q}_s^{\rho+1}).$$

On the other hand we have

(15′) $$F_{1\rho}(\omega;\xi) = \omega - (\phi_0 + \phi_1 + \cdots + \phi_\rho) \equiv 0(\mathfrak{p}_0^{\rho+1}).$$

Consequently, if we put

(16) $$G_\rho(\omega;\xi) = \prod_{j=1}^{h-1} F_{j\rho}(\omega;\xi),$$

then

$$G_\rho(\omega;\xi) \equiv 0(\mathfrak{p}_0^{\rho+1}\mathfrak{q}_1^{\rho+1} \cdots \mathfrak{q}_s^{\rho+1}),$$

or [12]

---

[12] If d_h was distinct from any one of the constants c_i, $i = 0, 1, \cdots, s$, then (15) would hold true for $i = 0, 1, \cdots, s$ and hence (16′) would follow directly, independently

$$(16') \qquad G_\rho(\omega;\xi) \equiv 0(\xi_1,\cdots,\xi_r)^{\rho+1}, \qquad (\rho = 0, 1, \cdots).$$

In view of (16′) it must be possible to express $G_\rho(\omega;\xi)$ as a form of degree $\rho+1$ in $\xi_1,\cdots,\xi_r$, with coefficients which are elements in $\mathfrak{o}$. Now any element in $\mathfrak{o}$ is an integral function of $\xi_1,\cdots,\xi_r$ and hence can be expressed in the form $g(\omega;\xi)/D(\xi)$, where g is a polynomial and D is the discriminant of $F(z;\xi)$. Hence we may write $G_\rho(\omega;\xi)$ in the form:

$$(17) \qquad G_\rho(\omega;\xi) = g_\rho(\omega;\xi)/D(\xi),$$

where g_ρ is a polynomial in ω, of degree $\leqq m-1$, with coefficients which are polynomials in $\xi_1,\cdots,\xi_r$ having no terms of degree $< \rho+1$. Now $G_\rho(\omega;\xi)$, according to (16), is of degree $m - m_h \leqq m-1$ in ω. Hence (17) must be an identity in $\omega, \xi_1,\cdots,\xi_r$. This is impossible, if ρ is sufficiently high. Namely, if the terms of lowest degree in D are of degree q, then the coefficient of the highest power of ω in $D\cdot G_\rho(\omega;\xi)$ begins with terms of degree q (since the leading coefficient of G_ρ is 1), while $g_\rho(\omega;\xi)$ has no terms of degree $<\rho+1$ in $\xi_1,\cdots,\xi_r$. Hence, if $\rho+1>q$, the relation (17) is impossible.

We have thus shown that there exist elements ω in $\mathfrak{o}$ with the property that, if $F(z;\xi) = N(z-\omega)$, then $F'_\omega(\omega;\xi) \not\equiv 0(\mathfrak{p}_0)$. Namely, we have also shown that if $\omega \equiv c_i(\mathfrak{p}_i)$, $i = 0, 1, \cdots, s$, then ω certainly enjoys this property, provided it is a primitive element over $K(\xi_1,\cdots,\xi_r)$ and provided $c_0 \neq c_i$, $i = 1, 2, \cdots, s$.

We now wish to prove that the first provision is a consequence of the second. Let ζ be any element in $\mathfrak{o}$, $\zeta \equiv b_i(\mathfrak{p}_i)$, $i = 0, 1, \cdots, s$, and let us assume that $b_0 \neq b_i$, $i = 1, 2, \cdots, s$. We have to show that ζ is a primitive element.

We fix in $\mathfrak{o}$ a primitive element ω such that $\omega \equiv c_i(\mathfrak{p}_i)$ and $c_0, c_1, \cdots, c_s$ are *distinct constants.* Let $N(z-\omega) = F(z;\xi)$ and let us again consider the factorization (11) of F in $K\{\xi_1,\cdots,\xi_r\}[z]$. Let $K^*\{\xi_1,\cdots,\xi_r\}$ denote the field of *meromorphic functions* of $\xi_1,\cdots,\xi_r$ and let Σ_i be the algebraic extension of this field defined by the irreducible equation $F_i(\omega;\xi) = 0$. Our field $\Sigma = K(\xi_1,\cdots,\xi_r,\omega)$ can be regarded as a subfield of Σ_i.[13] Let

---

of (15′), i. e. independently of the assumption that P is a simple point. The remainder of the proof is based on the impossibility of the congruences (16′). *Hence this also proves that any d_j must coincide with some c_i.*

[13] The following are well-known facts (see [7], p. 47). The ring of residual classes of $K^*\{\xi_1,\cdots,\xi_r\}[z]$ modulo F contains the field Σ and is the direct sum $\Sigma_1+\cdots+\Sigma_h$ of the fields Σ_i; each field Σ_i contains a subfield $\Sigma_i^{(0)} \cong \Sigma$, and the decomposition $a = a_1 + a_2 + \cdots + a_h$, $a \subset \Sigma$, $a_i \subset \Sigma_i^{(0)}$ sets up an isomorphism $a_i \to a_j$ between $\Sigma_i^{(0)}$

$N_i(z-\zeta)$ denote the norm of $z-\zeta$ over $K^*\{\xi_1,\cdots,\xi_r\}$, when ζ is regarded as an element of Σ_i. Since, in Σ, ζ is an integral function of $\xi_1,\cdots,\xi_r$, it follows that $N_i(z-\zeta)$ is a polynomial in z with coefficients in $\mathsf{K}\{\xi_1,\cdots,\xi_r\}$ and leading coefficient 1. Moreover, if $N(z-\zeta)$ denotes the usual norm of $z-\zeta$, when ζ is regarded as an element of Σ, then

$$N(z-\zeta)=\prod_{i=1}^{h} N_i(z-\zeta).$$

In particular, since $F_1(z;\zeta)=z-\phi$, $\phi\subset K\{\xi_1,\cdots,\xi_r\}$, we have $N_1(z-\zeta)=z-\psi$, where $\psi=\psi_0+\psi_1+\cdots$ (the expansion of ζ) and where $\psi_0=b_0$.

Now let us assume that ζ is not a primitive element. That $N(z-\zeta)$ is the power of a polynomial of degree $<m$, whence one of the factors $N_i(z-\zeta)$, $i\neq 1$, must be divisible by $z-\psi$. Let it be $N_2(z-\zeta)$. Then $N_2(z-\zeta)$ is necessarily a power of $z-\psi$:

$$N_2(z-\zeta)=(z-\psi)^{m_2}. \tag{18}$$

We consider in Σ_i the ring $\mathfrak{o}_i=\mathsf{K}\{\xi_1,\cdots,\xi_r\}\cdot\mathfrak{o}$ and the ideal $\mathfrak{p}_0^{(i)}=\mathfrak{o}_i\cdot\mathfrak{p}_0$. It is clear that $\mathfrak{p}_0^{(i)}$ contains the ideal of the non-units in $\mathsf{K}\{\xi_1,\cdots,\xi_r\}$ and that $\mathfrak{o}_i/\mathfrak{p}_0^{(i)}\cong\mathsf{K}$. Thus $\mathfrak{p}_0^{(i)}$ is prime and every element α in $\mathfrak{o}_i$ satisfies a congruence of the form $\alpha\equiv c(\mathfrak{p}_0^{(i)})$, $c\subset\mathsf{K}$. By (18), ζ, as an element of the field Σ_2, satisfies the relation $\zeta-\psi(\xi_1,\cdots,\xi_r)=0$, whence $\zeta\equiv\psi_0(\mathfrak{p}_0^{(2)})$, i. e. $\zeta\equiv b_0(\mathfrak{p}_0^{(2)})$, since $\psi_0=b_c$. We consider the contracted ideal $\mathfrak{p}_0^{(2)}\cap\mathfrak{o}$. This ideal contains the elements $\xi_1,\cdots,\xi_r$. Moreover, it also contains $\omega-d_2$, since the equation $F_2(\omega;\xi)=0$ yields the congruence $(\omega-d_2)^{m_2}\equiv 0(\mathfrak{p}_0^{(2)})$ (by (12$'$)). Now $\mathfrak{p}_0,\mathfrak{p}_1,\cdots,\mathfrak{p}_s$ are the only prime ideals in $\mathfrak{o}$ which divide the ideal $\mathfrak{o}\cdot(\xi_1,\cdots,\xi_r)$. Since $\mathfrak{p}_0^{(2)}\cap\mathfrak{o}$ is a prime 0-dimensional ideal, it must therefore coincide with one of the ideals $\mathfrak{p}_0,\cdots,\mathfrak{p}_s$. It cannot coincide with $\mathfrak{p}_0$, since $d_2\neq c_0$; therefore $\omega-d_2\not\equiv 0(\mathfrak{p}_0)$. Hence $\mathfrak{p}_0^{(2)}\cap\mathfrak{o}$ is one of the ideals $\mathfrak{p}_i$, $i\neq 0$. Let, say, $\mathfrak{p}_0^{(2)}\cap\mathfrak{o}=\mathfrak{p}_1$. We have then $\zeta\equiv b_0(\mathfrak{p}_1)$, and this is in contradiction with our hypothesis $b_0\neq b_i$, $i\neq 0$. Hence ζ is a primitive element, as was asserted.

To complete the proof of the theorem, we still have to show that, conversely, $F'_\omega(\omega;\xi_1,\cdots,\xi_r)\not\equiv 0(\mathfrak{p}_0)$ implies $c_0\neq c_i$, $i=1,2,\cdots,s$, where, again $F(z;\xi_1,\cdots,\xi_r)=N(z-\omega)$ and $\omega\equiv c_i(\mathfrak{p}_i)$. The hypothesis $F'_\omega(\omega;\xi)\not\equiv 0(\mathfrak{p}_0)$ signifies that c_0 is a simple root of $F(z;0,\cdots,0)$. Hence

---

and $\Sigma_j^{(0)}$. In particular we have $\omega=\omega_1+\cdots+\omega_h$, where $F_i(\omega_i;\xi_1,\cdots,\xi_r)=0$. There can be no confusion if, *for a fixed* i, we identify ω with ω_i and Σ with $\Sigma_i^{(0)}$.

the discriminant $D(\xi_1, \cdots, \xi_r)$ of $F(z;\xi)$ does not vanish, and consequently ω is a primitive element. Let ζ be an element in $\mathfrak{o}$ which assumes *distinct* values $b_0, b_1, \cdots, b_s$ at $\mathfrak{p}_0, \mathfrak{p}_1, \cdots, \mathfrak{p}_s$, and let, in our preceding notations, $N(z-\zeta) = G(z;\xi)$, $N_i(z-\zeta) = G_i(z;\xi)$. Here $G_1(z;\xi) = z - b_0 - \psi_1 - \psi_2 - \cdots$. Since $z - b_1$ is a factor of $G(z, 0, \cdots, 0)$, one of the polynomials $G_i(z; 0, \cdots, 0)$, $i \neq 1$, must be a power of $z - b_1$. Let, say, $G_2(z; 0, \cdots, 0) = (z - b_1)^{m_2}$. Then it follows that $\zeta \equiv b_1(\mathfrak{p}_0^{(2)})$ and we conclude as before that $\mathfrak{p}_0^{(2)} \circ \mathfrak{o} = \mathfrak{p}_1$. Now $\omega \equiv d_2(\mathfrak{p}_0^{(2)})$, hence $\omega \equiv d_2(\mathfrak{p}_1)$, i. e. $d_2 = c_1$. Since c_0 is a simple root of $F(z; 0, \cdots, 0)$, the remaining roots $d_2, \cdots, d_h$ are all distinct from d_1 $(= c_0)$. Consequently $c_1 \neq c_0$. Similarly $c_2 \neq c_0, \cdots, c_s \neq c_0$, and this completes the proof of our theorem.

6. *Second proof.* Let $\xi_i = \xi_i^{(1)}; \xi_i^{(2)}, \cdots, \xi_i^{(m)}$ be the conjugates of ξ_i over $\mathsf{K}(\xi_1, \cdots, \xi_r)$ $(i = r+1, \cdots, n)$ and let Σ^* be the Galois extension field obtained by adjoining to $\mathsf{K}(\xi_1, \cdots, \xi_r)$ the elements $\xi_{r+1}, \cdots, \xi_n$ and their conjugates. Let $\mathfrak{o}_j = \mathsf{K}[\xi_1, \cdots, \xi_r, \xi_{r+1}^{(j)}, \cdots, \xi_n^{(j)}]$ and let $\mathfrak{o}^*$ be the smallest ring in Σ^* which contains the m conjugate rings $\mathfrak{o}_1, \mathfrak{o}_2, \cdots, \mathfrak{o}_m$.[14]

By the isomorphism $\xi_i^{(1)} \to \xi_i^{(j)}$ between $\mathfrak{o}$ $(= \mathfrak{o}_1)$ and $\mathfrak{o}_j$, the ideal $\mathfrak{p}_0$ in $\mathfrak{o}$ is carried into a prime 0-dimensional ideal $\mathfrak{p}_0^{(j)}$ in $\mathfrak{o}_j$. Again the m ideals $\mathfrak{p}_0^{(j)}$, in general, need not be distinct. However, under the hypothesis of the theorem we prove that not only are they distinct, but that also *any two of the extended ideals* $\mathfrak{o}^*\mathfrak{p}_0^{(j)}$ *in* $\mathfrak{o}^*$ *have no common divisors.*

Let us consider, for instance, $\mathfrak{o}^*\mathfrak{p}_0^{(1)}$ and $\mathfrak{o}^*\mathfrak{p}_0^{(2)}$. These ideals are unmixed and 0-dimensional, since every element in $\mathfrak{o}^*$ is integrally dependent on $\mathfrak{o}_j$. Hence if they have a common divisor, they also have a common prime 0-dimensional divisor, say $\mathfrak{p}^*_0$. Let $\omega = \omega_1$ be a primitive element in $\mathfrak{o}$ over $\mathsf{K}(\xi_1, \cdots, \xi_r)$ and let ω_2 be its conjugate in $\mathfrak{o}_2$. By Theorem 2 we have $\omega_1 \equiv \psi_0 + \psi_1 + \cdots + \psi_{i-1}(\mathfrak{p}_0^{(1)i})$, where $\omega_1 = \psi_0 + \psi_1 + \cdots$ is the expansion of ω_1 into a power series of $\xi_1, \cdots, \xi_r$. Applying a substitution of the Galois group of $\Sigma^*/\mathsf{K}(\xi_1, \cdots, \xi_r)$ which carries $\mathfrak{o}_1$ into $\mathfrak{o}_2$, we get: $\omega_2 \equiv \psi_0 + \psi_1 + \cdots + \psi_{i-1}(\mathfrak{p}_0^{(2)i})$. Since $\mathfrak{p}^*_0$ is a common divisor of $\mathfrak{p}_0^{(1)}$ and $\mathfrak{p}_0^{(2)}$ it follows $\omega_1 \equiv \psi_0 + \cdots + \psi_{i-1}(\mathfrak{p}^{*i}_0)$ and $\omega_2 \equiv \psi_0 + \cdots + \psi_{i-1}(\mathfrak{p}^{*i}_0)$, whence

$$\omega_1 - \omega_2 \equiv 0(\mathfrak{p}^{*i}_0), \qquad (i = 1, 2, \cdots).$$

The validity of this congruence *for all* i implies $\omega_1 = \omega_2$, in contradiction with the hypothesis that ω_1 is a primitive element. This proves that $\mathfrak{o}^*\mathfrak{p}_0^{(1)}$ and $\mathfrak{o}^*\mathfrak{p}_0^{(2)}$ have no common divisors.

Let

$$\text{(19)} \qquad \mathfrak{o}^*\mathfrak{p}_0^{(j)} = \mathfrak{q}_1^{(j)} \cdots \mathfrak{q}_\sigma^{(j)} = [\mathfrak{q}_1^{(j)}, \cdots, \mathfrak{q}_\sigma^{(j)}]$$

[14] These m rings $\mathfrak{o}_j$ are not necessarily distinct.

be the decomposition of $\mathfrak{o}^*\mathfrak{p}_0^{(j)}$ into maximal primary (0-dimensional) components. Let $\mathfrak{q}_\alpha^{(j)}$ belong to the prime ideal $\mathfrak{p}_\alpha^{(j)}$. The prime ideals $\mathfrak{p}_1^{(j)}, \cdots, \mathfrak{p}_\sigma^{(j)}$ are the conjugates of the ideals $\mathfrak{p}_1^{(1)}, \cdots, \mathfrak{p}_\sigma^{(1)}$ respectively, under a substitution of the Galois group of Σ^* which carries $\mathfrak{o}_1$ into $\mathfrak{o}_j$. Since any two of the ideals $\mathfrak{o}^*\mathfrak{p}_0^{(j)}$ have no common divisors, *the σm ideals $\mathfrak{p}_\alpha^{(j)}$ are all distinct.*

Let now $\omega = \omega_1$ be an element of $\mathfrak{o}$ $(= \mathfrak{o}_1)$ such that $\omega \equiv c_0(\mathfrak{p}_0)$ and $\omega \not\equiv c_0(\mathfrak{p}_i)$, $i = 1, 2, \cdots, s$. Since $\omega \equiv c_0(\mathfrak{p}_0^{(1)})$, it follows from (19) that

$$(20) \qquad \omega - c_0 \equiv 0(\mathfrak{p}_\alpha^{(1)}), \qquad (\alpha = 1, 2, \cdots, \sigma).$$

We have $\mathfrak{o} \cdot (\xi_1, \cdots, \xi_r) \subseteqq \mathfrak{p}_0^{(1)}$, whence $\mathfrak{o}^*(\xi_1, \cdots, \xi_r) \subseteqq \mathfrak{p}_\alpha^{(1)}$, $\alpha = 1, 2, \cdots, \sigma$. Since the ideal $\mathfrak{o}^*(\xi_1, \cdots, \xi_r)$ is invariant under all the substitutions of the Galois group of Σ^*, it follows $\mathfrak{o}^*(\xi_1, \cdots, \xi_r) \subseteqq \mathfrak{p}_\alpha^{(j)}$, i. e. any of the $m\sigma$ ideals $\mathfrak{p}_\alpha^{(j)}$ is a divisor of $\mathfrak{o}^*(\xi_1, \cdots, \xi_r)$. Since $\mathfrak{p}_0, \mathfrak{p}_1, \cdots, \mathfrak{p}_s$ are the prime ideals of $\mathfrak{o} \cdot (\xi_1, \cdots, \xi_r)$, any $\mathfrak{p}_\alpha^{(j)}$ must be a divisor of one of these ideals. Now if $j \neq 1$, then $\mathfrak{p}_\alpha^{(j)}$ is not a divisor of $\mathfrak{p}_0$ $(= \mathfrak{p}_0^{(1)})$, since the prime ideals $\mathfrak{p}_\alpha^{(1)}$ of $\mathfrak{o}^*\mathfrak{p}_0^{(1)}$ are all distinct from any $\mathfrak{p}_\alpha^{(j)}$, $j \neq 1$. Hence, if $j \neq 1$, we must have, for some $\beta \neq 0$, $\mathfrak{p}_\beta \subseteqq \mathfrak{p}_\alpha^{(j)}$. Since, by assumption, $\omega_1 - c_0 \not\equiv 0(\mathfrak{p}_\beta)$ if $\beta \neq 0$, it follows

$$(21) \qquad \omega_1 - c_0 \not\equiv 0(\mathfrak{p}_\alpha^{(j)}), \qquad (j = 2, \cdots, m;\ \alpha = 1, 2, \cdots).$$

We now consider one of the ideals $\mathfrak{p}_\alpha^{(1)}$, say $\mathfrak{p}_1^{(1)}$. Let $\omega_1, \omega_2, \cdots, \omega_m$ be the conjugates of $\omega = \omega_1$; here $\omega_j \subset \mathfrak{o}_j$ and we do not yet assert that $\omega_1, \cdots, \omega_m$ are distinct, since we do not know whether ω is a primitive element. A substitution of the Galois group of Σ^* which carries $\mathfrak{o}_1$ into $\mathfrak{o}_2$, will carry ω_1 into ω_2 and also some ring $\mathfrak{o}_j$, $j \neq 1$, into $\mathfrak{o}_1$. Hence it will carry some ideal $\mathfrak{p}_\alpha^{(j)}$, $j \neq 1$, into $\mathfrak{p}_1^{(1)}$. Applying this substitution to (21) we obtain:

$$\omega_2 - c_0 \not\equiv 0(\mathfrak{p}_1^{(1)}).$$

In a similar manner we get $\omega_j - c_0 \not\equiv 0(\mathfrak{p}_1^{(1)})$, $j = 3, \cdots, m$. Reassuming and recalling (20), we have:

$$(22) \quad \omega_1 - c_0 \equiv 0(\mathfrak{p}_1^{(1)}),\ \omega_2 - c_0 \not\equiv 0(\mathfrak{p}_1^{(1)}), \cdots, \omega_m - c_0 \not\equiv 0(\mathfrak{p}_1^{(1)}).$$

From (22) it follows that $\omega_1 \neq \omega_j$, $j \neq 1$, whence ω *is a primitive element.* Moreover, let $N(z - \omega) = F(z; \xi_1, \cdots, \xi_r) = (z - \omega_1) \cdots (z - \omega_m)$, and let $\omega_j = c_j(\mathfrak{p}_1^{(1)})$, $j = 2, \cdots, m$. If we reduce the coefficients of $F(z; \xi_1, \cdots, \xi_r)$ mod $\mathfrak{p}_1^{(1)}$, we must replace $\xi_1, \cdots, \xi_r$ by zeros. Hence $F(z; 0, \cdots, 0) = (z - c_0)(z - c_2) \cdots (z - c_m)$. Since $c_j \neq c_0$, it follows that c_0 is a

simple root of $F(z; 0, \cdots, 0)$. Hence $F'_\omega(\omega; \xi_1, \cdots, \xi_r) \not\equiv 0(\mathfrak{p}_0)$, as was asserted.

Conversely, let us assume that c_0 is a simple root of $F(z; 0, \cdots, 0)$. This implies that $\omega_j \not\equiv c_0(\mathfrak{p}_\alpha^{(1)})$, $j \neq 1$, $\alpha = 1, 2, \cdots, \sigma$, since $\omega_1 \equiv c_0(\mathfrak{p}_\alpha^{(1)})$. Applying the substitutions of the Galois group of Σ^* which carry ω_j into ω_1, we get (21), and from this, retracing our reasoning which led to (21), we conclude that $\omega \not\equiv c_0(\mathfrak{p}_i)$, $i = 1, 2, \cdots, s$. Thus all the assertions of our theorem are proved.

7. It is clear that Theorem 4 remains true if we replace $\xi_1, \cdots, \xi_r$ by any other set of uniformizing parameters $\eta_1, \cdots, \eta_r$ in $\mathfrak{o}$, such that every element in $\mathfrak{o}$ is integrally dependent on $\eta_1, \cdots, \eta_r$. Let now $\eta_1, \cdots, \eta_r$ be any set of r algebraically independent elements in $\mathfrak{o}$, such that every element in $\mathfrak{o}$ is integrally dependent on $\eta_1, \cdots, \eta_r$. Let $F'_\omega(\omega; \eta_1, \cdots, \eta_r)$ be the different of an element ω in $\mathfrak{o}$, where $F(z; \eta_1, \cdots, \eta_r)$ is the norm $N(z - \omega)$ over $\mathsf{K}(\eta_1, \cdots, \eta_r)$. We consider the ideal $Z_{\eta_1, \ldots, \eta_r} = Z_{(\eta)}$ generated by the differents F'_ω as ω varies arbitrarily in $\mathfrak{o}$ (the h. c. d. of the ideals $\mathfrak{o} \cdot F'_\omega$, ω-arbitrary in $\mathfrak{o}$). Let $\mathfrak{p}_0$ be a prime zero-dimensional ideal in $\mathfrak{o}$, P—the corresponding point of V_r, and let $\eta_i \equiv c_i(\mathfrak{p}_0)$. If $\mathfrak{p}_0$ is an isolated component of the ideal $\mathfrak{o} \cdot (\eta_1 - c_1, \cdots, \eta_r - c_r)$, then P is a simple point of V_r and $\eta_1 - c_1, \cdots, \eta_r - c_r$ are uniformizing parameters for the whole neighborhood of P. It follows then, by Theorem 4, that for some ω in $\mathfrak{o}$ it is true that $F'_\omega \not\equiv 0(\mathfrak{p}_0)$. Hence $Z_{(\eta)} \not\equiv 0(\mathfrak{p}_0)$.

Conversely, let us assume that $Z_{(\eta)} \not\equiv 0(\mathfrak{p}_0)$. There exists then an element ω in $\mathfrak{o}$ for which $F'_\omega \not\equiv 0(\mathfrak{p}_0)$, i. e. such that $\omega \equiv d_0(\mathfrak{p}_0)$ and d_0 is a simple root of $F(z; c_1, \cdots, c_r)$. Let $\mathfrak{A} = \mathfrak{o} \cdot (\eta_1 - c_1, \cdots, \eta_r - c_r) = [\mathfrak{q}_0, \mathfrak{q}_1, \cdots, \mathfrak{q}_s]$, where $\mathfrak{q}_0, \mathfrak{q}_1, \cdots, \mathfrak{q}_s$ are all 0-dimensional and $\mathfrak{q}_0$ belongs to $\mathfrak{p}_0$. If we reduce the equation $F(\omega; \eta_1, \cdots, \eta_r) = 0$ modulo $\mathfrak{A}$, we get $F(\omega; c_1, \cdots, c_r) \equiv 0(\mathfrak{A})$ or

$$(23) \qquad (\omega - d_0) \prod_{i=1}^{m-1} (\omega - d_i) \equiv 0(\mathfrak{A}), \qquad d_i \neq d_0, \text{ if } i \neq 0,$$

where m is the relative degree of Σ over $\mathsf{K}(\eta_1, \cdots, \eta_r)$. *Since* $\omega - d_i \not\equiv 0(\mathfrak{p}_0)$, $i \neq 0$, *it follows by* (23), *that* $\omega - d_0 \equiv 0(\mathfrak{q}_0)$.

Let now ζ *be any element in* $\mathfrak{p}_0$. We form the norm $N(z - \bar{\zeta})$, where $\bar{\zeta} = t\zeta + \omega$, t—a parameter. This norm is a polynomial $G(z; t, \eta_1, \cdots, \eta_r)$ which for $t = 0$ becomes $F(z; \eta_1, \cdots, \eta_r)$. Moreover $G'_{\bar{\zeta}}(\bar{\zeta}; t, \eta_1, \cdots, \eta_r) = H(t)$, where $H(t)$ is a polynomial in t with coefficients in $\mathfrak{o}$. Now $H(0) = F'_\omega \not\equiv 0(\mathfrak{p}_0)$, hence the coefficients of $H(t)$ are not all $\equiv 0(\mathfrak{p}_0)$. As a consequence, if t_0 is a non-special value of t, and if $\bar{\zeta}_0 = t_0\zeta + \omega$, then

$G'_{0\bar{\zeta}_0} \not\equiv 0(\mathfrak{p}_0)$, where $G_0(z; \eta_1, \cdots, \eta_r) = N(z - \bar{\zeta}_0)$. Since $\bar{\zeta}_0 \equiv d_0(\mathfrak{p}_0)$, it follows, as above for ω, that $\bar{\zeta}_0 \equiv d_0(\mathfrak{q}_0)$, i. e. $t_0\zeta + \omega \equiv d_0(\mathfrak{q}_0)$. Since also $\omega \equiv d_0(\mathfrak{q}_0)$ and since we may assume $t_0 \neq 0$, we conclude that $\zeta \equiv 0(\mathfrak{q}_0)$. *Thus every element of* $\mathfrak{p}_0$ *is in* $\mathfrak{q}_0$, i. e. $\mathfrak{p}_0 = \mathfrak{q}_0$, whence $\mathfrak{p}_0$ *is an isolated component of the ideal* $\mathfrak{o} \cdot (\eta_1 - c_1, \cdots, \eta_r - c_r)$. We have therefore proved that $\mathfrak{p}_0$ *is an isolated component of* $\mathfrak{o} \cdot (\eta_1 - c_1, \cdots, \eta_r - c_r)$ *if and only if* $Z_{(\eta)} \not\equiv 0(\mathfrak{p}_0)$.

If we then denote by Z the h. c. d. of all ideal $Z_{(\eta)}$, as $\eta_1, \cdots, \eta_r$ vary arbitrarily in $\mathfrak{o}$ (*subject to the only condition that every element in* $\mathfrak{o}$ *be integrally dependent on* $\eta_1, \eta_2, \cdots, \eta_r$), we conclude with the following

THEOREM 5. *If P is a point on V_r and if $\mathfrak{p}_0$ is the corresponding prime 0-dimensional ideal in $\mathfrak{o}$, a necessary and sufficient condition that P be a simple point is that $Z \not\equiv 0(\mathfrak{p}_0)$.*

As a corollary we have that *the manifold of singular points of V_r is algebraic, of dimension $\leqq r - 1$.* It is the manifold given by the ideal Z.

8. Actually the singular manifold can be defined by an ideal which is possibly a multiple of Z and in whose construction intervene only the coördinates $\xi_1, \cdots, \xi_n$ of the general point of V_r. Namely, let P be a simple point of V_r, $\mathfrak{p}_0$—the corresponding ideal in $\mathfrak{o}$, and let $\xi_i \equiv c_i(\mathfrak{p}_0)$. We know from section 2 that if $\bar{\xi}_i = u_{i1}\xi_1 + \cdots + u_{in}\xi_n$, $i = 1, 2, \cdots, r$, and $\bar{\xi}_i \equiv d_i(\mathfrak{p}_0)$, then $\mathfrak{p}_0$ is an isolated component of the ideal $\mathfrak{o} \cdot (\bar{\xi}_1 - d_1, \cdots, \bar{\xi}_r - d_r)$, provided the coefficients u_{ij} are non-special. Moreover, every element in $\mathfrak{o}$ is then integrally dependent on $\bar{\xi}_1, \cdots, \bar{\xi}_r$. Hence P is a simple point if and only if $Z_{\bar{\xi}_1, \ldots, \bar{\xi}_r} \not\equiv 0(\mathfrak{p}_0)$, *for general values of the coefficients* u_{ij}. Now, assume that $Z_{\bar{\xi}_1, \ldots, \bar{\xi}_r} \not\equiv 0(\mathfrak{p}_0)$, i. e. there exists an element ω in $\mathfrak{o}$ such that $F'_\omega \not\equiv 0(\mathfrak{p}_0)$, where $F(z; \bar{\xi}_1, \cdots, \bar{\xi}_r) = N(z - \omega)$, the norm of $z - \omega$ relative to the field $\mathsf{K}(\bar{\xi}_1, \cdots, \bar{\xi}_r)$. We assert that *there also exists such an element* ω *which is linear and homogeneous in* $\xi_1, \cdots, \xi_n$. In fact, let

$$\mathfrak{o} \cdot (\bar{\xi}_1 - d_1, \cdots, \bar{\xi}_r - d_r) = [\mathfrak{p}_0, \mathfrak{q}_1, \cdots, \mathfrak{q}_s]$$

be the decomposition of $\mathfrak{o} \cdot (\bar{\xi}_1 - d_1, \cdots, \bar{\xi}_r - d_r)$ into primary (0-dimensional) components, and let $\xi_i \equiv c_{ij}(\mathfrak{p}_j)$, $i = 1, 2, \cdots, n$, $j = 0, 1, \cdots, s$, where $\mathfrak{q}_j$ belongs to the prime ideal $\mathfrak{p}_j$. Since $\mathfrak{p}_j$ and $\mathfrak{p}_0$ determine two distinct points of V_r, if $j \neq 0$, namely $P_j(c_{1j}, \cdots, c_{nj})$ and $P(c_{10}, \cdots, c_{n0})$, it follows that constants $\bar{v}_1, \cdots, \bar{v}_n$ can be found in such a manner that

$$\bar{v}_1 c_{1j} + \cdots + \bar{v}_n c_{nj} \neq \bar{v}_1 c_{10} + \cdots + \bar{v}_n c_{n0}, \qquad (j = 1, 2, \cdots, s).$$

Let $\bar{\xi} = \bar{v}_1\xi_1 + \cdots + \bar{v}_n\xi_n$. The value of the element $\bar{\xi}$ at P_0 is then distinct from its value at any of the points $P_1, \cdots, P_s$, and hence by Theorem 4 it follows that $F'_{\bar{\xi}}(\bar{\xi}; \bar{\xi}_1, \cdots, \bar{\xi}_r) \not\equiv 0(\mathfrak{p}_0)$, where $F(z; \bar{\xi}_1, \cdots, \bar{\xi}_r) = N(z - \bar{\xi})$. This proves our assertion.

From these considerations we conclude immediately with the following

THEOREM 6. *Let $\bar{\xi}_i = u_{i1}\xi_1 + \cdots + u_{in}\xi_n$, $i = 1, 2, \cdots, r+1$, be $r+1$ linear forms with indeterminate coefficients u_{ij}, and let $F(\bar{\xi}_1, \cdots, \bar{\xi}_{r+1}) = 0$ be the irreducible algebraic relation between the $\bar{\xi}_i$. Let $\mathfrak{Z}'$ be the ideal whose basis consists of the coefficients $f_\sigma(\xi_1, \cdots, \xi_n)$ of the various power products of the u_{ij} in the polynomial $F'_{\bar{\xi}_{r+1}}$ (the $\bar{\xi}$'s having been replaced in $F'_{\bar{\xi}_{r+1}}$ by the corresponding linear forms in the ξ's). The submanifold of V_r defined by this ideal $\mathfrak{Z}'$ is the manifold of singular points of V_r.*

9. Let $\eta_1, \cdots, \eta_r$ be elements in $\mathfrak{o}$ such that every element in $\mathfrak{o}$ is integrally dependent on $\mathsf{K}[\eta_1, \cdots, \eta_r]$ and let $a_1, \cdots, a_r$ be arbitrary constants. The ideal $\mathfrak{A} = \mathfrak{o} \cdot (\eta_1 - a_1, \cdots, \eta_r - a_r)$ is unmixed and zero dimensional,

$$\mathfrak{A} = [\mathfrak{q}_0, \mathfrak{q}_1, \cdots, \mathfrak{q}_s] = \mathfrak{q}_0\mathfrak{q}_1 \cdots \mathfrak{q}_s,$$

where $\mathfrak{q}_i$ is a primary ideal belonging to the zero-dimensional prime ideal $\mathfrak{p}_i$. Let ω be an element in $\mathfrak{o}$ and let $G(z; \eta_1, \cdots, \eta_r) = N(z - \omega)$ be the norm of $z - \omega$ with respect to the field $\mathsf{K}(\eta_1, \cdots, \eta_r)$. We have proved in sections 5, 6 that if $\mathfrak{q}_i \neq \mathfrak{p}_i$ and if $\omega \equiv c_i(\mathfrak{p}_i)$, then c_i is necessarily a multiple root of $G(z; a_1, \cdots, a_r)$. The following theorem gives a lower bound for the multiplicity of the root c_i:

THEOREM 7. *If $\mathfrak{q}_i$ belongs to the exponent ρ_i, then the multiplicity of the root c_i is $\geqq \rho_i$.*

Proof. Let $\omega_1, \cdots, \omega_\mu$ be an integral base of $\mathfrak{o}$ over the polynomial ring $\mathsf{K}[\eta_1, \cdots, \eta_r]$. We adjoin to the ground field K the indeterminates $u_1, u_2, \cdots, u_\mu$ and we consider the element $\omega_u = u_1\omega_1 + \cdots + u_\mu\omega_\mu$. Let

$$N(z - \omega_u) = F(z; \eta_1, \cdots, \eta_r; u_1, \cdots, u_\mu) = F(z; \eta; u)$$

be the norm of $z - \omega_u$ with respect to the field $\mathsf{K}(\eta; u)$. It is clear that if $\omega_i \equiv c_{ij}(\mathfrak{p}_j)$, $j = 0, 1, \cdots, s$, then $u_1c_{1j} + \cdots + u_\mu c_{\mu j}$ is a root of $F(z; a_1, \cdots, a_r; u_1, \cdots, u_\mu) (= F(z; a; u))$. It has been proved by van der Waerden [8] that, conversely, every root of $F(z; a; u)$ is a linear form $u_1d_1 + \cdots + u_\mu d_\mu$, $d_j \subset \mathsf{K}$, and that $(a_1, \cdots, a_r, d_1, \cdots, d_\mu)$ is a point of the variety whose general point has coördinates $\eta_1, \cdots, \eta_r, \omega_1, \cdots, \omega_\mu$. In

other words: $d_1, \cdots, d_\mu$ must be the values of $\omega_1, \cdots, \omega_\mu$ at one of the ideals $\mathfrak{p}_0, \mathfrak{p}_1, \cdots, \mathfrak{p}_s$. As a consequence, $F(z; a; u)$ is a product of factors $z - (u_1c_{1j} + \cdots + u_\mu c_{\mu j})$:

$$F(z; a; u) = \prod_{j=0}^{s} (z - c_{1j}u_1 - \cdots - c_{\mu j}u_\mu)^{\sigma_j}.$$

Since $\omega_1, \cdots, \omega_\mu$ is a base of $\mathfrak{o}$, it follows that

$$\mathfrak{p}_j = (\eta_1 - a_1, \cdots, \eta_r - a_r, \ \omega_1 - c_{1j}, \cdots, \omega_\mu - c_{\mu j}).$$

Hence for $j \neq j'$, *the factors* $z - c_{1j}u_1 - \cdots - c_{\mu j}u_\mu$ and $z - c_{1j'}u_1 - \cdots - c_{\mu j'}u_\mu$ *are distinct.*

We have $F(\omega_u; \eta; u) = 0$. Reducing the η modulo $\mathfrak{A}$ we get the congruence: $F(\omega_u; a; u) \equiv 0(\mathfrak{A})$, or

$$\prod_{j=0}^{s} (\omega_u - c_{1j}u_1 - \cdots - c_{\mu j}u_\mu)^{\sigma_j} \equiv 0(\mathfrak{A}). \tag{24}$$

This congruence should be intended in the following sense: if ω_u is replaced by $u_1\omega_1 + \cdots + u_\mu\omega_\mu$, then on the left-hand side we get a polynomial in $u_1, \cdots, u_\mu$, and the coefficients—element of $\mathfrak{o}$—of the various power products of the u's belong to $\mathfrak{A}$. Now, for a given j, we have

$$\omega_u - c_{1j}u_1 - \cdots - c_{\mu j}u_\mu = u_1(\omega_1 - c_{1j}) + \cdots + u_\mu(\omega_\mu - c_{\mu j}),$$

and the coefficients of this linear form in the u's do not all belong to $\mathfrak{p}_{j'}$, if $j' \neq j$. Hence, by (24),

$$[u_1(\omega_1 - c_{1j}) + \cdots + u_\mu(\omega_\mu - c_{\mu j})]^{\sigma_j} \equiv 0(\mathfrak{q}_j').$$

This implies that all the power products of degree σ_j in $\omega_1 - c_{1j}, \cdots, \omega_\mu - c_{\mu j}$ are in $\mathfrak{q}_j$. Since $\omega_1 - c_{1j}, \cdots, \omega_\mu - c_{\mu j}$ and the elements $\eta_1 - a_1, \cdots, \eta_r - a_r$ form a base of $\mathfrak{p}_j$, and since the elements $\eta_1 - a_1, \cdots, \eta_r - a_r$ are also in $\mathfrak{q}_j$, we conclude that $\mathfrak{p}_j^{\sigma_j} \equiv 0(\mathfrak{q}_j)$. *Hence* $\sigma_j \geqq \rho_j$.

Now let ω be any element in $\mathfrak{o}$,

$$\omega = P_1\omega_1 + \cdots + P_\mu\omega_\mu, \qquad P_i \subset K[\eta_1, \cdots, \eta_r],$$

and let $\omega \equiv c_j(\mathfrak{p}_j)$, $P_i \equiv P_{i0}(\mathfrak{A})$, $P_{i0} \subset K$. Then

$$N(z - \omega) = F(z; \eta_1, \cdots, \eta_r; \ P_1, \cdots, P_\mu) = G(z; \eta_1, \cdots, \eta_r),$$

whence

$$\begin{aligned} G(z; a_1, \cdots, a_r) &= F(z; a_1, \cdots, a_r; \ P_{10}, \cdots, P_{\mu 0}) \\ &= \prod_{j=0}^{s} (z - c_{1j}P_{10} - \cdots - c_{\mu j}P_{\mu 0})^{\sigma_j}. \end{aligned}$$

Since $c_{1j}P_{10} + \cdots + c_{\mu j}P_{\mu 0} = c_j$, our theorem follows.

Do there exist elements ω for which c_j is a root of multiplicity exactly ρ_j? The answer: not necessarily (except when $\rho_j = 1$, i. e. $\mathfrak{p}_j = \mathfrak{q}_j$, according to Theorem 4). Here is an example. Let $\eta^3{}_3 = \eta_1{}^2\eta_2$ be the defining equation of an algebraic surface and let $\eta_4 = \eta_3{}^2/\eta_1$. Consider the ring $\mathfrak{o} = \mathsf{K}[\eta_1, \eta_2, \eta_3, \eta_4]$. Every element of $\mathfrak{o}$ is integrally dependent on η_1, η_2, since $\eta_4{}^3 = \eta_1\eta_2{}^2$ (incidentally, it is not difficult to see that $\mathfrak{o}$ is integrally closed). The ideal $\mathfrak{A} = (\eta_1, \eta_2)$ is primary, $\mathfrak{A} = \mathfrak{q}$, and its prime ideal is $\mathfrak{p}_0 = (\eta_1, \eta_2, \eta_3, \eta_4)$. The exponent of $\mathfrak{A}$ is 2, since

$$\eta_3{}^2 = \eta_1\eta_4, \quad \eta_4{}^2 = \eta_2\eta_3, \quad \eta_3\eta_4 = \eta_1\eta_2.$$

On the other hand, the field $\mathsf{K}(\eta_1, \eta_2, \eta_3, \eta_4)$ is in the present case of relative degree 3 over $\mathsf{K}(\eta_1, \eta_2)$. Since $\mathfrak{A}$ itself is primary, it follows that if ω is any element of $\mathfrak{o}$ and if $F(z; \eta_1, \eta_2) = N(z - \omega)$, then $F(z, 0, 0)$ must have a triple root.

II. Properties of the conductor.

10. Let $\mathfrak{o}$, $\mathfrak{o}^*$ be two finite integral domains in Σ. We assume that $\mathfrak{o}$ is a subring of $\mathfrak{o}^*$ and that Σ is the quotient field of $\mathfrak{o}$. The conductor of $\mathfrak{o}$ with respect to $\mathfrak{o}^*$, in symbols $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*) = \mathfrak{c}$, is, by definition, the largest ideal in $\mathfrak{o}$ which is also an ideal in $\mathfrak{o}^*$. This implies that $\mathfrak{c}$ is the totality of all elements ξ in $\mathfrak{o}$ such that $\xi\mathfrak{o}^* \subset \mathfrak{o}$. Every element in $\mathfrak{o}^*$ can therefore be written in the form of a quotient η/ξ, $\eta \subset \mathfrak{o}$, and ξ—any element in $\mathfrak{c}$.

It is well known that $\mathfrak{c} \neq (0)$ *if and only if every element in* $\mathfrak{o}^*$ *is integrally dependent on* $\mathfrak{o}$. The proof is immediate. Namely, assume $\mathfrak{c} \neq (0)$ and let ξ be an element in $\mathfrak{c}$, different from zero. The elements in $\mathfrak{o}$ are integrally dependent on the ring of polynomials $\mathsf{P} = \mathsf{K}[\eta_1, \cdots, \eta_r]$, where $\eta_1, \cdots, \eta_r$ are suitable elements in $\mathfrak{o}$. Hence $\mathfrak{o}$ is a finite P-module (since the totality of all elements of Σ which are integrally dependent on P is also a finite P-module; see [7], p. 94). Let $\omega_1, \cdots, \omega_m$ be a P-basis for $\mathfrak{o}$. Then $\mathfrak{o}^*$ is contained in the finite P-module $(\omega_1/\xi, \cdots, \omega_m/\xi)$ and hence $\mathfrak{o}^*$ itself is a finite P-module. Consequently, every element in $\mathfrak{o}^*$ is integrally dependent on P.

Conversely, assume that every element in $\mathfrak{o}^*$ is integrally dependent on $\mathfrak{o}$, whence also on P (in view of the transitivity of integral dependence). Then $\mathfrak{o}^*$ is a finite P-module. Let $\omega^*_1, \cdots, \omega^*_m$ be a P-basis for $\mathfrak{o}^*$. We can write each ω^*_i in the form ω_i/ξ, where ω_i, $\xi \subset \mathfrak{o}$, ξ being a common denominator (since $\mathfrak{o}^*$ is contained in the quotient ring of $\mathfrak{o}$). Hence $\xi\mathfrak{o}^* \subset \mathfrak{o}$, $\xi \subset \mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*)$, $\xi \neq 0$, q. e. d.

Let $\mathfrak{p}_0$ be a prime 0-dimensional ideal in $\mathfrak{o}$. We assume that $\mathfrak{c}(\mathfrak{o},\mathfrak{o}^*)\neq(0)$ and we consider the extended ideal $\mathfrak{o}^*\mathfrak{p}_0$. Since the elements in $\mathfrak{o}^*$ are integrally dependent on $\mathfrak{o}$, the ideal $\mathfrak{o}^*\mathfrak{p}_0$ is unmixed and zero-dimensional.

THEOREM 8. *If* $\mathfrak{c}(\mathfrak{o},\mathfrak{o}^*)\not\equiv 0(\mathfrak{p}_0)$, *then* $\mathfrak{o}^*\mathfrak{p}_0$ *is a prime ideal in* $\mathfrak{o}^*$.

Proof. Let ξ be an element in $\mathfrak{c}$, not in $\mathfrak{p}_0$, and let $\xi\equiv d(\mathfrak{p}_0)$, $d\neq 0$, $d\subset\mathsf{K}$. For any element ω^* in $\mathfrak{o}^*$ we have a relation of the form $\xi\omega^*=\omega\subset\mathfrak{o}$. Reducing this relation modulo $\mathfrak{o}^*\mathfrak{p}_0$ we find $d\omega^*\equiv c(\mathfrak{o}^*\mathfrak{p}_0)$, where $\omega\equiv c(\mathfrak{p}_0)$. Since $d\neq 0$, we conclude that the ring of residual classes $\mathfrak{o}^*/\mathfrak{o}^*\mathfrak{p}_0$ is a field simply isomorphic to K. Hence $\mathfrak{o}^*\mathfrak{p}_0$ is prime (and zero-dimensional).

11. In this and in the following sections we shall derive several properties of the conductor in the special case when $\mathfrak{o}^*$ is the integral closure of $\mathsf{P}=\mathsf{K}[\eta_1,\cdots,\eta_r]$ and $\mathfrak{o}=\mathsf{K}[\eta_1,\cdots,\eta_r,\omega]$. Here $\eta_1,\cdots,\eta_r$ are algebraically independent elements of Σ and ω ($\subset\mathfrak{o}^*$) is a primitive element of Σ with respect to the field $\mathsf{K}(\eta_1,\cdots,\eta_r)$. Let n be the relative degree $[\Sigma\colon\mathsf{K}(\eta_1,\cdots,\eta_r)]$, and let $\mathfrak{c}(\omega)$ be the conductor $\mathfrak{c}(\mathfrak{o};\mathfrak{o}^*)$.

Let $\bar{\mathfrak{p}}$ be an arbitrary zero-dimensional ideal in P. We may assume that $\bar{\mathfrak{p}}=\mathsf{P}\cdot(\eta_1,\cdots,\eta_r)$.

THEOREM 9. *A necessary and sufficient condition that* $1,\omega,\cdots,\omega^{n-1}$ *form a* K*-basis of the* K*-module* $\mathfrak{o}^*/\mathfrak{o}^*\bar{\mathfrak{p}}$, *is that the contracted ideal* $\mathfrak{c}(\omega)\frown\mathsf{P}$ *should not be divisible by* $\bar{\mathfrak{p}}$.

Proof. That the condition is sufficient is trivial. In fact, assuming $\mathfrak{c}(\omega)\frown\mathsf{P}\not\equiv 0(\bar{\mathfrak{p}})$, let g be an element in $\mathfrak{c}(\omega)\frown\mathsf{P}$ but not in $\bar{\mathfrak{p}}$. Let ζ be an arbitrary element in $\mathfrak{o}^*$ and let

$$g\zeta=g_0+g_1\omega+\cdots+g_{n-1}\omega^{n-1},\qquad g_i\subset\mathsf{P}.$$

If we reduce $g,g_0,\cdots,g_{n-1}$ modulo $\bar{\mathfrak{p}}$, we get, in view of $g\not\equiv 0(\bar{\mathfrak{p}})$:

$$\zeta\equiv c_0+c_1\omega+\cdots+c_{n-1}\omega^{n-1}(\mathfrak{o}^*\bar{\mathfrak{p}}),\tag{25}$$

as was asserted.

Somewhat more difficult is the proof that the condition is necessary. We shall make use of the following well-known relation between $\mathfrak{c}(\omega)$, the complementary module e of $\mathfrak{o}^*$ and the different G'_ω of ω:

$$\mathfrak{c}(\omega)=\mathsf{e}G'_\omega.^{15}\tag{26}$$

---

[15] As was pointed out by Schmeidler [6], this relation holds true for fields of algebraic functions of several variables, the proof being the same as in the case of algebraic functions of one variable.

We assume then that for every element ζ in $\mathfrak{o}^*$ a congruence such as (25) holds true. We can write (25) as follows:

$$(25') \qquad \zeta = c_0 + c_1\omega + \cdots + c_{n-1}\omega^{n-1} + \eta_1 A_1 + \cdots + \eta_r A_r, \quad A_i \subset \mathfrak{o}^*.$$

If we apply the congruence (25), taking as ζ any of the elements $A_1, \cdots, A_r$, we derive from (25′) a congruence of the form:

$$\zeta \equiv f_1(\omega; \eta_1, \cdots, \eta_r)(\mathfrak{o}^*\bar{\mathfrak{p}}^2),$$

where $f_1 = f_{10} + f_{11}\omega + \cdots + f_{1,n-1}\omega^{n-1}$, and f_{1i} is a polynomial of first degree in $\eta_1, \cdots, \eta_r$. Applying repeatedly this procedure we get more generally the following congruence:

$$(27) \qquad \zeta \equiv f_{\rho-1}(\omega; \eta_1, \cdots, \eta_r)(\mathfrak{o}^*\bar{\mathfrak{p}}^\rho),$$

where ρ is an arbitrary integer and $f_{\rho-1}$ is a polynomial, of degree $\leqq n-1$ in ω and of degree $\leqq \rho - 1$ in $\eta_1, \cdots, \eta_r$.

We now consider the complementary module $\mathfrak{e}'$ of $\mathfrak{o}$, i. e. the set of all elements η in Σ such that $T(\eta\xi)$ (trace of $\eta\xi$ with respect to the field $\mathsf{K}(\eta_1, \cdots, \eta_r)$) is in P if ξ is in $\mathfrak{o}$. It is well known that the elements

$$1/G'_\omega, \omega/G'_\omega, \cdots, \omega^{n-1}/G'_\omega,$$

form a P-basis for $\mathfrak{e}'$. Here G'_ω is the different of ω; $G(\eta_1, \cdots, \eta_r; \omega) = 0$ is the irreducible equation for ω.

In view of the existence of a finite P-basis in $\mathfrak{o}^*$ and in $\mathfrak{e}'$, it follows that the traces of *all* the products $\eta\zeta$, $\eta \subset \mathfrak{e}'$, $\zeta \subset \mathfrak{o}^*$, can be written as rational functions in $\eta_1, \cdots, \eta_r$ with the same denominator. Let $h(\eta) = h(\eta_1, \cdots, \eta_r)$ be this common denominator. Let us fix an element ζ in $\mathfrak{o}^*$ and an element η in $\mathfrak{e}'$ and let $T(\zeta\eta) = g(\eta)/h(\eta)$. We apply the congruence (27). The trace of the product $f_{\rho-1}\cdot\eta$ is a polynomial $\psi(\eta)$ in $\eta_1, \cdots, \eta_r$, since $\eta \subset \mathfrak{e}'$. The trace of the product $(\zeta - f_{\rho-1})\cdot\eta$ will be, by (27), of the form $g_\rho(\eta)/h(\eta)$, where $g_\rho \equiv 0(\bar{\mathfrak{p}}^\rho)$. We have then the relation: $g/h = \psi + g_\rho/h$, or $g = h\psi + g_\rho$, or finally the congruence

$$(28) \qquad g(\eta) \equiv 0(h(\eta), \bar{\mathfrak{p}}^\rho).$$

We separate in the polynomial $h(\eta)$ the factors which belong to $\bar{\mathfrak{p}}$ (i. e. those which vanish for $\eta_1 = \cdots = \eta_r = 0$) from the remaining factors. We write then $h(\eta) = \sigma(\eta)h_1(\eta)$, where $h_1(0, \cdots, 0) \neq 0$ and $\sigma(\eta)$ is the product of the factors which are $\equiv 0(\bar{\mathfrak{p}})$. If such factors are not present in h, then we put $\sigma(\eta) = 1$. We proceed to show that from the fact that the congruence (28)

holds true for any positive integer ρ, it necessarily follows that *$g(\eta)$ is divisible by $\sigma(\eta)$*.

It is clear that in (28) it is permissible to divide through g and h by any common factor of g and h, provided—if that common factor begins with terms of lowest degree ρ_0—that we replace ρ by $\rho - \rho_0$. Hence, to prove our assertion it is sufficient to show that *g and h cannot be relatively prime unless $\sigma(\eta) = 1$.*

Performing, if necessary, a preliminary linear homogeneous transformation on the elements $\eta_1, \cdots, \eta_r$, we may assume that $h(\eta)$, considered as polynomial in η_r, has leading coefficient 1. Since $g = h\psi + g_\rho$, we deduce then that the resultant R of g and h (considered as polynomials in η_r) coincides with the resultant of g_ρ and h. Let

$$h = \eta_r^{\nu} + A_1(\eta_1, \cdots, \eta_{r-1})\eta_r^{\nu-1} + \cdots + A_\nu(\eta_1, \cdots, \eta_{r-1}),$$

$$g_\rho = B_0(\eta_1, \cdots, \eta_{r-1})\eta_r^{\mu} + \cdots + B_\mu(\eta_1, \cdots, \eta_{r-1}).$$

Since $g_\rho \equiv 0(\bar{\mathfrak{p}}^\rho)$, it follows that $B_\mu, B_{\mu-1}, \cdots, B_{\mu-\rho+1}$ begin with terms of lowest degree not less than $\rho, \rho - 1, \cdots, 1$ respectively. Now let us assume that $\sigma(\eta) \neq 1$. Then A_ν begins with terms of lowest degree $\geqq 1$. We apply a theorem on the resultant proved in [9], p. 250. If we attach to A_ν the weight 1, to $B_\mu, B_{\mu-1}, \cdots, B_{\mu-\rho+1}$ the weights $\rho, \rho - 1, \cdots, 1$ respectively, and to the remaining coefficients the weight zero, then—as a special case of the theorem referred to—every term in the resultant R is of weight $\geqq \rho$. It follows that R, as a polynomial in $\eta_1, \cdots, \eta_{r-1}$, begins with terms of lowest degree $\geqq \rho$. Since ρ is an arbitrary integer, it follows that R is identically zero. Hence h and g have a common factor, q. e. d.[16]

---

[16] Our assertion can also be proved without making use of the properties of the resultant. Assume that g and h are relatively prime. Then the intersection of the ideals $(h, \bar{\mathfrak{p}}^\rho)$, $\rho = 1, 2, \cdots$ is at most $(r-2)$-dimensional, since both g and h belong to it. If then $\bar{\eta}_1, \ldots, \bar{\eta}_{r-2}$ are linear forms in $\eta_1, \cdots, \eta_r$ with non special coefficients, then the intersection of the ideals $(h, \bar{\eta}_1, \cdots, \bar{\eta}_{r-2}, \bar{\mathfrak{p}}^\rho)$ is at most zero-dimensional. We may assume that $\bar{\eta}_1, \cdots, \bar{\eta}_{r-2}$ coincide with $\eta_1, \cdots, \eta_{r-2}$ respectively, and we denote by $\mathfrak{B}$ the intersection of the ideals $(h, \eta_1, \cdots, \eta_{r-2}, \bar{\mathfrak{p}}^\rho)$, $\rho = 1, 2, \ldots$. $\mathfrak{B}$ is at most 0-dimensional. If $\sigma \neq 1$, then $h \equiv 0(\bar{\mathfrak{p}})$ and hence $\mathfrak{B} \equiv 0(\bar{\mathfrak{p}})$. Consequently, $\mathfrak{B}$ is 0-dimensional and one of its prime ideals is $\bar{\mathfrak{p}}$. If q is the exponent of the corresponding primary component of $\mathfrak{B}$, then (see [7], p. 49) $\bar{\mathfrak{p}}^q \equiv 0(\mathfrak{B}, \bar{\mathfrak{p}}^{q+1}) \equiv 0(h, \eta_1, \cdots, \eta_{r-2}, \bar{\mathfrak{p}}^\rho, \bar{\mathfrak{p}}^{q+1})$. Since ρ is arbitrary, this implies $\bar{\mathfrak{p}}^q \equiv 0(h, \eta_1, \cdots, \eta_{r-2}, \bar{\mathfrak{p}}^{q+1})$. It is immediately seen that this congruence leads to the following absurd conclusion: if $h = h_j + h_{j+1} + \cdots$, where h_j is a form of degree j, and if $h'_j = h'_j(\eta_{r-1}, \eta_r)$ denotes the sum of terms in h_j which depend only on η_{r-1} and η_r, then every form in η_{r-1}, η_r, of degree q, is divisible by h'_j (or is identically zero, if $h'_j = 0$).

One can also observe that the congruence (28) implies that the hypersurface $g = 0$ has at the origin a contact of infinite order with every algebraic curve which passes

Since h is the least common denominator of all the traces $T(\zeta\eta)$, $\zeta \subset \mathfrak{o}^*$, $\eta \subset \mathfrak{e}'$, *we conclude by the above result that necessarily* $h \not\equiv 0(\bar{\mathfrak{p}})$. Now if $T(\zeta\eta) = g/h$, then $T(\zeta \cdot \eta h) = g$. From this it follows $h\eta \subset \mathfrak{e}$, η arbitrary in $\mathfrak{e}'$. In particular $h \cdot (1/G'_\omega) \subset \mathfrak{e}$, i. e. $h \subset \mathfrak{e}G'_\omega$. Hence, by (26), $h \equiv 0(\mathfrak{c}(\omega))$, and since $h \not\equiv 0(\bar{\mathfrak{p}})$, our theorem is proved.

12. We are now in position to prove the converse of Theorem 8, always under the assumption: $\mathfrak{o} = K[\eta_1, \cdots, \eta_r, \omega]$, $\mathfrak{o}^*$-integral closure of $\mathfrak{o}$.

THEOREM 10. *If $\mathfrak{p}_0$ is a prime zero-dimensional ideal in $\mathfrak{o}$ and if $\mathfrak{o}^*\mathfrak{p}_0$ is prime in $\mathfrak{o}^*$, then* $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*) \not\equiv 0(\mathfrak{p}_0)$.

Proof. We may assume that $\eta_i \equiv 0(\mathfrak{p}_0)$, $i = 1, 2, \cdots, r$, and also that $\omega \equiv 0(\mathfrak{p}_0)$. Then $\mathfrak{p}_0 = (\eta_1, \eta_2, \cdots, \eta_r, \omega)$. Let V_r be the variety given in S_{r+1} by the defining equation $G(\eta_1, \cdots, \eta_r, \omega) = 0$ and let A be the point $(0, 0, \cdots, 0)$ of V_r which corresponds to $\mathfrak{p}_0$. Let M be the subvariety of V_r defined by the conductor $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*)$. A general line on A will intersect V_r, outside of A, in points which are not on M. Subjecting the elements $\eta_1, \cdots, \eta_r, \omega$ to a preliminary linear homogeneous transformation, we may assume that the line $\eta_1 = \cdots = \eta_r = 0$ is general in the sense just specified. Let $[\mathfrak{q}_0, \mathfrak{q}_1, \cdots, \mathfrak{q}_s]$ be decompositions of the ideal $\mathfrak{o}(\eta_1, \cdots, \eta_r)$ into 0-dimensional primary components, $\mathfrak{p}_0, \mathfrak{p}_1, \cdots, \mathfrak{p}_s$—the corresponding prime ideals. Our assumption concerning the line $\eta_1 = \cdots = \eta_r = 0$ implies that $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*) \not\equiv 0(\mathfrak{p}_i)$, $i = 1, 2, \cdots, s$. Let $\omega \equiv c_i(\mathfrak{p}_i)$. The constants $c_0 = 0$, $c_1, \cdots, c_s$ are distinct, since $\mathfrak{p}_i = (\eta_1, \cdots, \eta_r, \omega - c_i)$. By Theorem 8 it follows that the ideals

$$\mathfrak{p}^*_i = \mathfrak{o}^*\mathfrak{p}_i = \mathfrak{o}^*(\eta_1, \cdots, \eta_r, \omega - c_i), \qquad (i = 1, 2, \cdots, s)$$

are prime. By our assumption, the ideal

$$\mathfrak{p}^*_0 = \mathfrak{o}^*\mathfrak{p}_0 = \mathfrak{o}^*(\eta_1, \cdots, \eta_r, \omega)$$

is also prime. From this it follows:

$$(29) \quad [\mathfrak{p}^*_0, \mathfrak{p}^*_1, \cdots, \mathfrak{p}^*_s] = \mathfrak{p}^*_0\mathfrak{p}^*_1 \cdots \mathfrak{p}^*_s \equiv \mathfrak{o}^*(\eta_1, \cdots, \eta_r, \prod_{i=0}^{s} (\omega - c_i)).$$

Let now ζ be an arbitrary element in $\mathfrak{o}^*$ and let $\zeta \equiv d_i(\mathfrak{p}^*_i)$, $i = 0, 1, \cdots s$. We can find a polynomial $f(\omega)$, with constant coefficients (of degree $\leqq s$ in ω), such that $f(c_i) = d_i$. For such a polynomial $f(\omega)$ we will have $\zeta - f(\omega) \equiv 0([\mathfrak{p}^*_0, \cdots, \mathfrak{p}^*_s])$, or, in view of (29),

$$\zeta = f(\omega) + \alpha^*_1\eta_1 + \cdots + \alpha^*_r\eta_r + \beta^* \cdot \prod_{i=0}^{s} (\omega - c_i), \quad \alpha^*_i, \beta^* \subset \mathfrak{o}^*.$$

through the origin and lies on the hypersurface $h = 0$. Hence every such irreducible algebraic curve must also lie on the hypersurface $g = 0$, which shows that the two hypersurfaces $g = 0$ and $h = 0$ must have a common component.

Applying to the elements α^*_i and β^* the considerations just applied to ζ, we find that ζ satisfies a congruence of the form

$$\zeta \equiv f_1(\omega; \eta_1, \cdots, \eta_r)(\mathfrak{p}^*{}_0{}^2\mathfrak{p}^*{}_1{}^2 \cdots \mathfrak{p}^*{}_s{}^2),$$

where f_1 is a polynomial (of degree $\leqq 1$ in $\eta_1, \cdots, \eta_r$). More generally, we find

$$\zeta \equiv f_\rho(\omega; \eta_1, \cdots, \eta_r)(\mathfrak{p}^*{}_0{}^{\rho+1} \cdots \mathfrak{p}^*{}_s{}^{\rho+1}),$$

where f_ρ is a polynomial in $\omega, \eta_1, \cdots, \eta_r$, whose degree in ω may be assumed to be $\leqq n-1$. Now if $\rho+1$ is sufficiently high, then $\mathfrak{p}^*{}_i{}^{\rho+1} \equiv 0(\mathfrak{o}^*\mathfrak{q}_i)$.[17] Hence, for ρ sufficiently high, we will have

$$\zeta \equiv f_\rho(\omega; \eta_1, \cdots, \eta_r)(\mathfrak{o}^*(\eta_1, \cdots, \eta_r)),$$

or, letting $f_\rho(\omega; 0, \cdots, 0) = c_0 + c_1\omega + \cdots + c_{n-1}\omega^{n-1}$, $c_i \subset \mathsf{K}$,

$$\zeta \equiv c_0 + c_1\omega + \cdots + c_{n-1}\omega^{n-1}(\mathfrak{o}^*\bar{\mathfrak{p}}).$$

Such a congruence holds true for any element ζ in $\mathfrak{o}^*$. Hence, by Theorem 9, we must have $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*) \frown \mathsf{P} \not\equiv 0(\bar{\mathfrak{p}})$. This implies $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*) \not\equiv 0(\mathfrak{p}_i)$, $i = 0, 1, \cdots, s$, in particular $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*) \not\equiv 0(\mathfrak{p}_0)$, which proves our assertion.

13. In the preceding section we have characterized in $\mathfrak{o}$ the prime zero-dimensional divisors $\mathfrak{p}_0$ of the conductor $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*)$, by means of the decomposition of the extended ideal $\mathfrak{o}^*\mathfrak{p}_0$. Namely, $\mathfrak{p}_0$ is a divisor of $\mathfrak{c}(\mathfrak{o}, \mathfrak{o}^*)$ if and only if the zero-dimensional ideal $\mathfrak{o}^*\mathfrak{p}_0$ is not prime in $\mathfrak{o}^*$. This is a relative characterization. It is of interest to give an intrinsic characterization in $\mathfrak{o}^*$ of those zero-dimensional ideals $\mathfrak{p}^*{}_0$ which are extended ideals $\mathfrak{o}^*\mathfrak{p}_0$ of ideals $\mathfrak{p}_0$ in some subrings of $\mathfrak{o}^*$, such as $\mathfrak{o}$. The question is then the following. We consider subrings $\mathfrak{o}$ of $\mathfrak{o}^*$ which satisfy the following conditions: (1) $\mathfrak{o}^*$ is the integral closure of $\mathfrak{o}$; (2) $\mathfrak{o}$ is generated by $r+1$ elements, $\mathfrak{o} = \mathsf{K}[\eta_1, \cdots, \eta_{r+1}]$; (3) $\mathfrak{o}$ and $\mathfrak{o}^*$ have the same quotient field. Given a prime zero-dimensional ideal $\mathfrak{p}^*{}_0$ in $\mathfrak{o}^*$, we ask: under what condition does there exist a subring $\mathfrak{o}$ such that $\mathfrak{p}^*{}_0 = \mathfrak{o}^*\mathfrak{p}_0$, where $\mathfrak{p}_0 = \mathfrak{p}^*{}_0 \frown \mathfrak{o}$? We proceed to prove that such a subring $\mathfrak{o}$ exists if and only if *the rank of the* K*-module* $\mathfrak{p}^*{}_0/\mathfrak{p}^*{}_0{}^2$ *is not greater than* $r+1$.[18]

Proof. Consider, quite generally, any zero-dimensional ideal $\mathfrak{p}^*{}_0$ in $\mathfrak{o}^*$, and let ρ be the rank of the ring $\mathfrak{p}^*{}_0/\mathfrak{p}^*{}_0{}^2$, considered as a K-module. Let

---

[17] Since $\mathfrak{o}^*\mathfrak{p}_i = \mathfrak{p}_i{}^*$, the ideal $\mathfrak{o}^*\mathfrak{q}_i$ is primary and belongs to the ideal $\mathfrak{p}_i{}^*$. Namely, in the first place $\mathfrak{o}^*\mathfrak{q}_i$ is zero-dimensional. Let $\mathfrak{p}'^*$ be one of its prime ideals, and let $\mathfrak{p}'^* \frown \mathfrak{o} = \mathfrak{p}'$. Also $\mathfrak{p}'$ is prime and zero-dimensional. Since $\mathfrak{p}'^* \supseteqq \mathfrak{q}_i$, we have $\mathfrak{p}' \supseteqq \mathfrak{q}_i$, whence $\mathfrak{p}' = \mathfrak{p}_i$. Consequently $\mathfrak{p}'^* \supseteqq \mathfrak{o}^*\mathfrak{p}' = \mathfrak{o}^*\mathfrak{p}_i = \mathfrak{p}^*{}_i$, i. e., $\mathfrak{p}'^* = \mathfrak{p}^*{}_i$.

[18] Hence, in particular, if the point which corresponds to $\mathfrak{p}_0{}^*$ on the variety determined by the ring $\mathfrak{o}^*$, is simple, because then the rank is r (see Theorem 3.2).

$\omega_1, \cdots, \omega_\rho$ be a K-basis of $\mathfrak{p}^*_0/\mathfrak{p}^{*2}_0$. We have evidently: $(\omega_1, \cdots, \omega_\rho, \mathfrak{p}^{*2}_0) = \mathfrak{p}^*_0$. From this we deduce (see section 2) that $\mathfrak{p}^*_0$ is an isolated component of the ideal $\mathfrak{o}^*(\omega_1, \cdots, \omega_\rho)$. Since $\mathfrak{p}^*_0$ is zero-dimensional, we must have $\rho \geqq r$ (see footnote 3).

Now let us assume that the rank ρ is $\leqq r+1$. Let $\mathfrak{o}^* = \mathsf{K}[\xi_1, \cdots, \xi_n]$ and let us assume, as usual, that $\mathfrak{p}^*_0 = (\xi_1, \cdots, \xi_n)$. We can then take as a K-basis for $\mathfrak{p}^*_0/\mathfrak{p}^{*2}_0$ ρ linear forms in the ξ's. We may assume that $\xi_1, \cdots, \xi_\rho$ form such a K-basis. Let $\xi_i \equiv \sum_{j=1}^{\rho} c_{ij}\xi_j(\mathfrak{p}^{*2}_0)$, $i = 1, 2, \cdots, n$. The rank of the matrix (c_{ij}) is ρ. Given ρ forms $\eta_j = \sum_{i=1}^{n} u_{ji}\xi_i$, they will form a K-basis for $\mathfrak{p}^*_0/\mathfrak{p}^{*2}_0$ if and only if the determinant of the ρ by ρ matrix $(u_{ji}) \cdot (c_{i\nu})$ is $\neq 0$. We consider the two possible cases: $\rho = r$, $\rho = r+1$.

Let $\rho = r$. We choose the coefficients u_{ji} so that the above determinant be $\neq 0$ and that, in addition, every element of $\mathfrak{o}^*$ be integrally dependent on $\eta_1, \cdots, \eta_r$. Then $\mathfrak{p}^*_0$ is an isolated component of the *zero-dimensional* ideal $\mathfrak{o}^*(\eta_1, \cdots, \eta_r)$. Let $\mathfrak{p}^*_1, \cdots, \mathfrak{p}^*_s$ be the other prime ideals of $\mathfrak{o}^*(\eta_1, \cdots, \eta_r)$. We choose in $\mathfrak{o}^*$ a primitive element ω (with respect to the field $\mathsf{K}(\eta_1, \cdots, \eta_r)$) such that $\omega \equiv 0(\mathfrak{p}^*_0), \omega \not\equiv 0(\mathfrak{p}^*_i)$, $i = 1, 2, \cdots, s$. Then $\mathfrak{p}^*_0 = \mathfrak{o}^*(\eta_1, \cdots, \eta_r, \omega)$ and the ring $\mathfrak{o} = \mathsf{K}[\eta_1, \cdots, \eta_r, \omega]$ satisfied all our conditions.

Let $\rho = r+1$. We first choose the coefficients of the first r rows of the matrix $(u_{ji}) \cdot (c_{i\nu})$ in such a manner that their matrix be of rank r and that every element in $\mathfrak{o}^*$ be integrally dependent on $\eta_1, \cdots, \eta_r$. With this choice of the first r rows we are certain that when $u_{r+1,1}, \cdots, u_{r+1,n}$ have non-special values, then the determinant of the $(r+1)$-row matrix $(u_{ji})(c_{i\nu})$ is $\neq 0$. Let $[\mathfrak{q}^*_0, \mathfrak{q}^*_1, \cdots, \mathfrak{q}^*_s]$ be the decomposition of the *zero-dimensional* ideal $\mathfrak{o}^*(\eta_1, \cdots, \eta_r)$ into primary components. Let $\mathfrak{p}^*_0, \mathfrak{p}^*_1, \cdots, \mathfrak{p}^*_s$ be the corresponding prime ideals. Again, for non-special values of the coefficients $u_{r+1,i}$, the element $\eta_{r+1} = \sum_{i=1}^{n} u_{r+1,i}\xi_i$ will satisfy the conditions: $\eta_{r+1} \not\equiv 0(\mathfrak{p}_i)$, $i = 1, 2, \cdots, s$ (since the n elements ξ_j are not all $\equiv 0(\mathfrak{p}_i)$, if $i \neq 0$). Finally, for non-special coefficients $u_{r+1,i}$ the element η_{r+1} will be a primitive element of $\mathfrak{o}^*$ with respect to the field $\mathsf{K}(\eta_1, \cdots, \eta_r)$. If we choose the coefficients $u_{r+1,i}$ so as to satisfy all these conditions, it will follow that $\mathfrak{o}^*(\eta_1, \cdots, \eta_r, \eta_{r+1}) = \mathfrak{p}^*_0$ and that the ring $\mathfrak{o} = \mathsf{K}[\eta_1, \cdots, \eta_{r+1}]$ satisfies all our requirements, q. e. d.[19]

[19] The example at the end of section 9 illustrates the possibility $\rho > r+1$. In that example we have $r = 2$, $\mathfrak{p}_0{}^* = (\eta_1, \eta_2, \eta_3, \eta_4)$, and it is easily seen that the four elements η_i are linearly independent modulo $\mathfrak{p}_0{}^{*2}$. Hence $\rho = 4$.

III. Normal varieties in the affine space.

14. In part I we have given several characterizations of simple and multiple points of a V_r in S_n, mostly of ideal theoretic nature. We have seen that these characterizations are not so much properties of the set of coördinates $\xi_1, \cdots, \xi_n$ of the general point of V_r, as of the whole ring $\mathfrak{o} = K[\xi_1, \cdots, \xi_n]$, namely of the zero-dimensional ideals in $\mathfrak{o}$. If $\eta_1, \cdots, \eta_m$ is another set of elements in $\mathfrak{o}$ such that $\mathfrak{o} = K[\eta_1, \cdots, \eta_m]$, then the η's are the coördinates of the general point of a variety W_r in an S_m, birationally equivalent to V_r. The points of V_r and W_r are in $(1,1)$ correspondence without exceptions (everything being confined to the points in the affine space, i. e. at finite distance), since their points are in $(1,1)$ correspondence with the prime zero-dimensional ideals in $\mathfrak{o}$. To simple points of V_r correspond simple points of W_r, and conversely. Topologically speaking, V_r and W_r (points at infinity excluded) are homeomorphic loci ("open" or "relative" circuits). We shall say that V_r and W_r are *integrally equivalent*, alluding to the fact that they are both defined in terms of one and the same integral domain $\mathfrak{o}$ in the field Σ. Thus $\mathfrak{o}$ determines a *class* of birationally equivalent varieties any two of which are integrally equivalent.

We shall say that an algebraic variety V_r in an affine S_n is *normal in the affine space*, if the coördinates $\xi_1, \cdots, \xi_n$ of the general point of V_r give rise to an integral domain $\mathfrak{o} = K[\xi_1, \cdots, \xi_n]$ which is *integrally closed in its quotient field*. In the sequel we will speak of *the* normal variety determined by an integrally closed finite integral domain $\mathfrak{o}$ in the field Σ, meaning by this any one of the integrally equivalent varieties determined by $\mathfrak{o}$.

15. We prove now the following

LEMMA. *Let V_r be an irreducible r-dimensional algebraic variety in S_n, $\mathfrak{o} = K[\xi_1, \cdots, \xi_n]$ the corresponding integral domain. If V_ρ is a ρ-dimensional irreducible algebraic subvariety of V_r and $\mathfrak{p}_\rho$ is the corresponding ρ-dimensional prime ideal in $\mathfrak{o}$, then a sufficient condition that V_ρ be a non-singular manifold of V_r* (i. e. *that not all points of V_ρ be multiple points of V_r*) *is that there should exist $r - \rho$ elements $\eta_1, \cdots, \eta_{r-\rho}$ in $\mathfrak{o}$ such that $\mathfrak{p}_\rho$ is an isolated component of the ideal $\mathfrak{o} \cdot (\eta_1, \cdots, \eta_{r-\rho})$.*

This lemma generalizes the property which we have used in section 2 for the definition of a simple point $(\rho = 0)$.

For the proof we consider the ring of residual classes $\bar{\mathfrak{o}} = \mathfrak{o}/\mathfrak{p}_\rho$. If $\bar{\xi}_1, \cdots, \bar{\xi}_n$ denote the elements of $\bar{\mathfrak{o}}$ which correspond to $\xi_1, \cdots, \xi_n$ respectively in the homomorphism $\mathfrak{o} \cong \bar{\mathfrak{o}}$, then it is clear that $\bar{\xi}_1, \cdots, \bar{\xi}_n$ are the coördinates of the general point of V_ρ. We take a point P of V_ρ which does

not lie on either one of the following two subvarieties of V_ρ (both of dimension $\leqq \rho - 1$): (1) the variety of singular points of V_ρ (the points of V_ρ which are multiple *for* V_ρ); (2) the variety in which V_ρ intersects the possible other components of the variety V' defined by the ideal $\mathfrak{o}\cdot(\eta_1,\cdots,\eta_{r-\rho})$ (since $\mathfrak{p}_\rho$ is an isolated component of $\mathfrak{o}\cdot(\eta_1,\cdots,\eta_{r-\rho})$, V_ρ itself is one of the irreducible components of V'). Let $\bar{\mathfrak{p}}_0$ and $\mathfrak{p}_0$ be the prime 0-dimensional ideals in $\bar{\mathfrak{o}}$ and $\mathfrak{o}$ respectively which correspond to the point P, regarded either as a point of V_ρ or as a point of V_r. Clearly $\mathfrak{p}_\rho \equiv 0(\mathfrak{p}_0)$ and $\mathfrak{p}_0 \rightarrow \bar{\mathfrak{p}}_0$ in the homomorphism between $\mathfrak{o}$ and $\bar{\mathfrak{o}}$. Since P is a simple point for V_ρ, it follows that the rank of the K-module $\bar{\mathfrak{p}}_0/\bar{\mathfrak{p}}_0^2$ is ρ (Theorem 3.2). Let $\bar{\eta}_{r-\rho+1},\cdots,\bar{\eta}_r$ be a K-module basis of $\bar{\mathfrak{p}}_0/\bar{\mathfrak{p}}_0^2$, and let $\eta_{r-\rho+1},\cdots,\eta_r$ be any set of ρ elements in $\mathfrak{o}$ belonging to the residual classes $\bar{\eta}_{r-\rho+1},\cdots,\bar{\eta}_r$ respectively. Any element $\bar{\omega}$ in $\bar{\mathfrak{o}}$ satisfies a congruence of the form $\bar{\omega} \equiv d_0 + \sum_{i=1}^{\rho} d_i\bar{\eta}_{r-\rho+i}(\bar{\mathfrak{p}}_0^2)$, $d_i \subset K$. Since the largest ideal in $\mathfrak{o}$ which is mapped upon $\bar{\mathfrak{p}}_0^2$ in the homomorphism $\mathfrak{o} \simeq \bar{\mathfrak{o}}$ is the ideal $(\mathfrak{p}_0^2, \mathfrak{p}_\rho)$ it follows that any element ω of $\mathfrak{o}$ satisfies a congruence of the form

$$\omega \equiv d_0 + \sum_{i=1}^{\rho} d_i\eta_{r-\rho+i}(\mathfrak{p}_0^2, \mathfrak{p}_\rho). \tag{30}$$

Now, in view of our hypothesis concerning the ideal $\mathfrak{o}\cdot(\eta_1,\cdots,\eta_{r-\rho})$ and the point P, it follows that $\mathfrak{p}_0$ does not divide any of the primary components of $\mathfrak{o}\cdot(\eta_1,\cdots,\eta_{r-\rho})$ distinct from $\mathfrak{p}_\rho$. Hence there exists an element α in $\mathfrak{o}$ such that $\alpha \not\equiv 0(\mathfrak{p}_0)$ and $\alpha\mathfrak{p}_\rho \equiv 0(\eta_1,\cdots,\eta_{r-\rho})$. Hence for any element ζ in $\mathfrak{p}_\rho$ we have: $\alpha\zeta = A_1\eta_1 + \cdots + A_{r-\rho}\eta_{r-\rho}$, where $A_i \subset \mathfrak{o}$. If then $\alpha \equiv c_0(\mathfrak{p}_0)$, $c_0 \neq 0$ and $A_j \equiv c_j(\mathfrak{p}_0)$, $c_0,\cdots,c_{r-\rho} \subset K$, it follows immediately, since the elements $\zeta,\eta_1,\cdots,\eta_{r-\rho}$ are in $\mathfrak{p}_0$, that a congruence of the form

$$c_0\zeta \equiv c_1\eta_1 + \cdots + c_{r-\rho}\eta_{r-\rho}(\mathfrak{p}_0^2),$$

holds true for any element ζ in $\mathfrak{p}_\rho$. From this, in view of (30), we conclude the r elements $\eta_1,\cdots,\eta_r$ form a K-module basis for $\mathfrak{p}_0/\mathfrak{p}_0^2$, which is then of rank r. Hence P is a simple point of V_r. Thus it is proved that V_ρ contains points which are simple for V_r, q. e. d.

The above lemma implies as an immediate consequence the following important property of normal varieties:

THEOREM 11. *The singular manifold of a normal variety V_r in an affine space is of dimension $\leqq r - 2$.*

Proof. It is sufficient to prove that every $(r-1)$-dimensional subvariety V_{r-1} of V_r is non-singular. Let $\mathfrak{p}_{r-1}$ be the prime ideal of V_{r-1} in $\mathfrak{o}$. Since

$\mathfrak{o}$ is integrally closed, the primary ideals belonging to the *minimal* ideal $\mathfrak{p}_{r-1}$ are the symbolic powers $\mathfrak{p}^{(i)}_{r-1}$ of $\mathfrak{p}_{r-1}$, $i = 1, 2, \cdots$.[20] Let η be an element in $\mathfrak{p}_{r-1}$ but not in $\mathfrak{p}^{(2)}_{r-1}$. Then, by the principal ideal theorem in the integrally closed ring $\mathfrak{o}$, it follows that $\mathfrak{p}_{r-1}$ is an isolated component of $\mathfrak{o} \cdot \eta$. Our theorem now follows from the above lemma in the case $\rho = r - 1$.

16. If V_r is not normal, whence the corresponding ring $\mathfrak{o} = K[\xi_1, \cdots, \xi_n]$ is not integrally closed, the passage to the integral closure $\mathfrak{o}^*$ of $\mathfrak{o}$ defines a birational transformation of V_r into the normal variety V^*_r defined by $\mathfrak{o}^*$ (or into any other variety V'_r integrally equivalent to V^*_r). We shall say that V^*_r is *the derived normal variety* of V_r. For lack of a better word we shall use the term "integral closure" to denote the birational transformation which carries V_r into V^*_r. By Theorem 11, we may say that *the effect of the integral closure is the elimination of all singular manifolds of dimension* $r - 1$. Thus for algebraic curves, *the integral closure transformation resolves all the singularities of the curve* (singularities of finite distance).[21] For an algebraic surface the integral closure resolves the multiple curves (at finite distance). *A normal surface in the affine space has only a finite number of singularities* (at finite distance).

One more remark concerning the birational correspondence between V_r and V^*_r. If P is a point of V_r, $\mathfrak{p}_0$—the corresponding 0-dimensional prime ideal in $\mathfrak{o}$, then $\mathfrak{o}^*\mathfrak{p}_0$ is also 0-dimensional and unmixed. If $\mathfrak{o}^*\mathfrak{p}_0 = [\mathfrak{q}^*_1, \cdots, \mathfrak{q}^*_s]$, where each $\mathfrak{q}^*_i$ belongs to a prime 0-dimensional ideal $\mathfrak{p}^*_i$ in $\mathfrak{o}^*$, and if P^*_i is the point on V^*_r determined by $\mathfrak{p}^*_i$, then to the point P of V_r there correspond on V_r the points $P^*_1, \cdots, P^*_s$. On the other hand, every prime 0-dimensional ideal $\mathfrak{p}^*_0$ in $\mathfrak{o}^*$ determines uniquely its contracted ideal $\mathfrak{p}_0 = \mathfrak{o} \cap \mathfrak{p}^*_0$ in $\mathfrak{o}$. Hence to a point on V^*_r there corresponds a unique point on V_r. By Theorem 8, if $\mathfrak{p}_0$ does not divide the conductor of $\mathfrak{o}$ relative to $\mathfrak{o}^*$, then $s = 1$ and $\mathfrak{o}^*\mathfrak{p}_0$ is prime in $\mathfrak{o}^*$. Now assume that P is a simple point of V_r, and let $\xi_1, \cdots, \xi_r$ be uniformizing parameters for the whole neighborhood of P, where we assume that $\xi_i \equiv 0(\mathfrak{p}_0)$ and that every element of $\mathfrak{o}$ (whence also of $\mathfrak{o}^*$) is integrally dependent on $\xi_1, \cdots, \xi_r$. By Theorem 4, there exists an element ω in $\mathfrak{o}$ such that $F'_\omega(\omega; \xi_1, \cdots, \xi_r) \not\equiv 0(\mathfrak{p}_0)$, where $F(z; \xi_1, \cdots, \xi_r)$ is the norm of $z - \omega$ over $K(\xi_1, \cdots, \xi_r)$. Since $\mathfrak{p}_0 = \mathfrak{p}^*_i \cap \mathfrak{o}$, it follows that $F'_\omega(\omega; \xi_1, \cdots, \xi_r) \not\equiv 0(\mathfrak{p}^*_i)$, and hence, by Theorem 5, we conclude that P^*_i *is a simple point of* V^*_r and that $\mathfrak{p}^*_i$ is an isolated component of the ideal $\mathfrak{o}^*(\xi_1, \cdots, \xi_r)$. Since $\mathfrak{o}^*\mathfrak{p}_0$ divides this last ideal, it follows

[20] See van der Waerden [7], p. 105.

[21] See Muhly and Zariski [3].

a fortiori that each $\mathfrak{p}^*_i$ is an isolated component of $\mathfrak{o}^*\mathfrak{p}_0$, whence $\mathfrak{q}^*_i = \mathfrak{p}^*_i$ and $\mathfrak{o}^*\mathfrak{p}_0 = [\mathfrak{p}^*_1, \cdots, \mathfrak{p}^*_s]$. Now at each point P^*_i we have a definite isomorphic mapping τ_i of $\mathfrak{o}^*$ upon a subring H_i of the ring of formal power series of $\xi_1, \cdots, \xi_r$. *These s mappings τ_i must coincide on* $\mathfrak{o}$. In fact, since $\xi_1, \cdots, \xi_r$ are uniformizing parameters for the whole neighborhood of P on V_r, we have a mapping τ of $\mathfrak{o}$ upon a subring H of $K\{\xi_1, \cdots, \xi_r\}$. Let ω be any element of $\mathfrak{o}$ and let

$$\tau\omega = \phi_0 + \phi_1(\xi_1, \cdots, \xi_r) + \phi_2(\xi_1, \cdots, \xi_r) + \cdots \subset H.$$

Then $\omega \equiv \phi_0 + \phi_1 + \cdots + \phi_m (\mathfrak{p}_0^{m+1})$, m arbitrary, and consequently also $\omega \equiv \phi_0 + \phi_1 + \cdots + \phi_m (\mathfrak{p}^*{}_0{}^{m+1})$. If $\tau_i\omega \rightarrow \phi_0{}^{(i)} + \phi_1{}^{(i)} + \phi_2{}^{(i)} + \cdots \subset H_i$, then also $\omega \equiv \phi_0{}^{(i)} + \phi_1{}^{(i)} + \cdots + \phi_m{}^{(i)} (\mathfrak{p}^*{}_0{}^{m+1})$. In view of the uniqueness of the polynomial $\phi_0{}^{(i)} + \cdots + \phi_m{}^{(i)}$, it follows $\phi_m{}^{(i)} = \phi_m$, and this proves our assertion $\tau = \tau_i$ on $\mathfrak{o}$, $i = 1, 2, \cdots, s$. Since $\mathfrak{o}$ and $\mathfrak{o}^*$ have the same quotient field Σ, it follows that these mappings also coincide on $\mathfrak{o}^*$, *whence necessarily* $s = 1$.

Reassuming, we have the following

THEOREM 12. *If V^*_r is the derived normal variety of V_r (in the affine space), then to each point of V^*_r there corresponds a unique point of V_r, while to every point P of V_r there correspond at most a finite number of points of V^*_r. This number can be greater than one only if P is a singular point of V_r and lies on the subvariety defined on V_r by the conductor of* $\mathfrak{o}$ *relative to* $\mathfrak{o}^*$.[22]

This theorem shows that the birational transformation between V_r and V^*_r is free from fundamental points on either variety.

IV. Normal varieties in the projective space.

17. A normal variety in the affine space may have singularities at infinity. Concerning these, Theorem 11 gives us no information whatever. It may very well happen that a normal V_r has a singular V_{r-1} at infinity. Hence, from a projective—and consequently also from an algebro-geometric point of view—Theorem 11 is not significant. Also the relationship between a V_r and its derived normal variety V^*_r in an affine space has no invariantive character from an algebro-geometric standpoint. Thus, for instance, the birational

[22] The preceding proof shows that every point of the variety defined by the conductor is necessarily a singular point of V_r. The converse is of course not always true. On the other hand, it should be noted that to P there may correspond a unique point of $V_r{}^*$ even if P lies on the variety of the conductor. Namely, we may have $\mathfrak{o}^*\mathfrak{p}_0 = \mathfrak{q}^*$, $\mathfrak{q}^*$ primary (not prime).

correspondence between V_r and V^*_r may have fundamental loci at infinity. We shall therefore now deal with the extension of the notion of a normal variety to projective spaces.

Let $y_0, y_1, \cdots, y_n$ be homogeneous coördinates in an n-dimensional projective space P_n. The quotients $x_i = y_i/y_0$, $i = 1, 2, \cdots, n$, are point coördinates in the affine S_n consisting of those points of P_n which are not on the hyperplane $y_0 = 0$. Let V_r be an irreducible r-dimensional variety in P_n, *not in the hyperplane* $y_0 = 0$, and let $\xi_1, \cdots, \xi_n$ be the coördinates of the general point of V_r in S_n. The field $\Sigma = \mathsf{K}(\xi_1, \cdots, \xi_n)$ is of degree of transcendency r over K. The variety V_r is defined in P_n by the H-ideal (*homogeneous ideal*) $\mathfrak{P}$ in $\mathsf{K}[y_0, \cdots, y_n]$, generated by all forms $f(y_0, \cdots, y_n)$ such that $f(1, \xi_1, \cdots, \xi_n) = 0$.

Now suppose that we choose as hyperplane at infinity another hyperplane, say $c_0y_0 + c_1y_1 + \cdots + c_ny_n = 0$, which does not contain the variety V_r, and let S'_n be the affine space consisting of the points of P_n which are not on this hyperplane. If, say, $c_i \neq 0$, then the n quotients $y_j/(c_0y_0 + \cdots + c_ny_n)$, $j = 0, 1, \cdots, i-1, i+1, \cdots, n$, can be taken as coördinates in S'_n. The coördinates of the general point of V_r, considered in S'_n, are then

$$\xi'_0 = 1/(c_0 + c_1\xi_1 + \cdots + c_n\xi_n), \quad \xi'_j = \xi_j/(c_0 + c_1\xi_1 + \cdots + c_n\xi_n), \quad j \neq i.$$

The ring of polynomials in the ξ'_j contains also the quotient

$$\xi'_i = \xi_i/(c_0 + c_1\xi_1 + \cdots + c_n\xi_n),$$

since $c_0\xi'_0 + \cdots + c_n\xi'_n = 1$ and $c_i \neq 0$. Thus, inasmuch as we regard V_r as a variety in the affine space S'_n, it determines in Σ the integral domain

$$\text{(31)} \quad \mathfrak{o}' = \mathsf{K}[1/(c_0 + c_1\xi_1 + \cdots + c_n\xi_n), \\ \xi_1/(c_0 + c_1\xi_1 + \cdots + c_n\xi_n), \cdots, \xi_n/(c_0 + c_1\xi_1 + \cdots + c_n\xi_n)].$$

We go back to the homogeneous coördinates. If we regard $y_0, \cdots, y_n$ as non-homogeneous coördinates in an affine S_{n+1}, then the homogeneous ideal $\mathfrak{P}$ defines an $(r+1)$-dimensional irreducible variety W_{r+1} in S_{n+1}. The coördinates $\xi^*_0, \xi^*_1, \cdots, \xi^*_n$ of the general point of W_{r+1} are the residual classes in $\mathsf{K}[y]/\mathfrak{P}$ containing $y_0, y_1, \cdots, y_n$ respectively. If we imagine P_n as being the hyperplane at infinity of S_{n+1}, then W_{r+1} is the hypercone which projects V_r from the origin $y_0 = \cdots = y_n = 0$.

Let $\Sigma^* = \mathsf{K}(\xi^*_0, \xi^*_1, \cdots, \xi^*_n)$ be the field of rational functions on W_{r+1}. Its subfield $\mathsf{K}(\xi^*_1/\xi^*_0, \cdots, \xi^*_n/\xi^*_0)$ is simply isomorphic with the field $\Sigma = \mathsf{K}(\xi_1, \cdots, \xi_n)$, and in this isomorphism to ξ^*_i/ξ^*_0 there corresponds the element ξ_i. This follows immediately from the relationship between the homogeneous ideal $\mathfrak{P}$ in $\mathsf{K}[y]$ and the defining ideal of V_r in $\mathsf{K}[x]$. If then we identify ξ^*_i/ξ^*_0 with ξ_i, we may regard Σ as a subfield of Σ^*. The degree of

transcendency of Σ^* is one greater than that of Σ. Since $\xi^*_i = \xi_i \xi^*_0$ (in view of our identification), we have $\Sigma^* = \Sigma(\xi^*_0)$. Hence Σ^* is a *simple transcendental extension* of Σ.[23] Thus we have an invariantive relationship between Σ and Σ^*, independent on the particular projective model V_r of the field Σ. We note the existence of the group of relative automorphisms τ of Σ^* with respect to Σ; these are all of the form:

$$(32) \qquad \tau: \quad \xi_i \to \xi_i; \; \xi^*_0 \to \frac{\alpha\xi^*_0 + \beta}{\gamma\xi^*_0 + \delta}, \quad \begin{matrix} \alpha, \beta, \gamma, \delta \subset \Sigma, \\ \alpha\delta - \beta\gamma \neq 0. \end{matrix}$$

Such a relative automorphism is described by a birational transformation of the hypercone W_{r+1} into itself which leaves invariant each generator.

Consider the special automorphisms τ_t: $\xi_i \to \xi_i$; $\xi^*_0 \to t\xi^*_0$, $t \subset \mathsf{K}$. We shall say that an element ω^* in Σ^* is *homogeneous of degree* ν, if $\tau_t(\omega^*) = t^\nu \omega^*$, t arbitrary in K. *Any homogeneous element* ω^* *of degree* ν *belonging to the ring* $\mathfrak{o}^* = \mathsf{K}[\xi^*_0, \cdots, \xi^*_n]$ *is a form of degree* ν *in* $\xi^*_0, \cdots, \xi^*_n$. In fact, let $\omega^* = f_\rho(\xi^*) + f_{\rho+1}(\xi^*) + \cdots + f_\sigma(\xi^*)$, where $f_i(\xi^*)$ is a form of degree i in $\xi^*_0, \cdots, \xi^*_n$ and $f_\rho \neq 0$. We must have:

$$\begin{aligned} \tau_t(\omega^*) &= t^\rho f_\rho(\xi^*) + t^{\rho+1} f_{\rho+1}(\xi^*) + \cdots + t^\sigma f_\sigma(\xi^*) \\ &= t^\nu f_\rho(\xi^*) + t^\nu f_{\rho+1}(\xi^*) + \cdots + t^\nu f_\sigma(\xi^*), \quad t\text{—arbitrary.} \end{aligned}$$

This implies $\rho = \nu$, $f_{\rho+1} = \cdots = f_\sigma = 0$.

The elements of Σ are homogeneous of degree 0. Conversely, *any element of* Σ^* *which is homogeneous of degree zero, is an element of* Σ. Namely, it is clear that the elements of $\Sigma^* = \Sigma(\xi^*_0)$ which are left invariant by all the relative automorphisms $\xi^*_0 \to t\xi^*_0$ are necessarily elements of Σ.

In terms of the field Σ^* we are now in position to define quite generally the notion of the *homogeneous coördinates of the general point of an algebraic variety* V_r *in the projective space* P_n:

Any set of $n + 1$ elements in Σ^*, say $\zeta^*_0, \zeta^*_1, \cdots, \zeta^*_n$, are homogeneous coördinates of the general point of V_r in P_n if:

(a) $\zeta^*_i / \zeta^*_0 = \xi_i$;

(b) The field $\mathsf{K}(\zeta^*_0, \zeta^*_1, \cdots, \zeta^*_n)$ is a transcendental extension of Σ.

The condition (b) implies that each element ζ^*_i is transcendental with respect to Σ. A particular set of homogeneous coördinates is given by the elements $\xi^*_0, \xi^*_1, \cdots, \xi^*_n$. The most general set of homogeneous coördinates $\zeta^*_0, \zeta^*_1, \cdots, \zeta^*_n$ (relative to our fixed coördinate system $y_0, \cdots, y_n$ in P_n) is obtained by multiplying the coördinates ξ^*_i by any element σ^* in Σ^*, pro-

---

[23] That ξ_0^* is transcendental with respect to Σ $(= \mathsf{K}(\xi_1^*/\xi_0^*, \cdots, \xi_n^*/\xi_0^*)$, follows also from the fact that all the algebraic relations between $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$ are homogeneous (or consequences of homogeneous relations).

vided, however, that $\sigma^*\xi^*_i$ be transcendental over Σ. Thus, we may also say that any set $\zeta^*_0, \cdots, \zeta^*_n$ of homogeneous coördinates of the general point of V_r is of the form: $\rho^*, \rho^*\xi_1, \cdots, \rho^*\xi_n$, ρ^* an element of Σ^* which is transcendental over Σ. An immediate consequence of this is the following: *the correspondence* $\xi^*_i \to \zeta^*_i$ *sets up an isomorphism between the two rings* $K[\xi^*]$ and $K[\zeta^*]$.

18. Definition. *Let* $\xi^*_0, \xi^*_1, \cdots, \xi^*_n$ *(or* $\sigma^*\xi^*_0, \cdots, \sigma^*\xi^*_n, \sigma^* \subset \Sigma^*$*) be homogeneous coördinates of the general point of a* V_r *in* P_n. *The variety* V_r *shall be said to be normal (in* P_n*), if the ring* $\mathfrak{o}^* = K[\xi^*_0, \cdots, \xi^*_n]$ *(or the ring* $K[\sigma^*\xi^*_0, \cdots, \sigma^*\xi^*_n]$*) is integrally closed in its quotient field.*

It is clear that the above defining property of a normal variety V_r is independent of the choice of the factor of proportionality σ^*, in view of the remark at the end of the preceding section.

When we speak in the sequel of a normal variety, it will be understood that the variety is normal in the projective space.

Let us take any given hyperplane $c_0y_0 + \cdots + c_ny_n = 0$ in P_n as hyperplane at infinity, and let S'_n be the corresponding affine space. We assert that *if* V_r *is normal, then it is also normal in the affine space* S'_n (provided, of course, that V_r does not lie entirely in the preassigned hyperplane at infinity). We have to show that the ring (31), or what is the same, that the ring

$$\mathfrak{o}' = K[\xi^*_0/\sum_{i=0}^{n} c_i\xi^*_i, \xi^*_1/\sum_{i=0}^{n} c_i\xi^*_i, \cdots, \xi^*_n/\sum_{i=0}^{n} c_i\xi^*_i]$$

is integrally closed in Σ. Let $\sum_{i=0}^{n} c_i\xi^*_i = \eta^*$, and let ω be an element in Σ which depends integrally on $\mathfrak{o}'$. Then it is clear that there exists an integer ρ such that $\omega \cdot (\eta^*)^\rho$ depends integrally on $\mathfrak{o}^* = K[\xi^*_0, \cdots, \xi^*_n]$. Hence $\omega \cdot (\eta^*)^\rho \subset \mathfrak{o}^*$,

$$\omega \cdot (\eta^*)^\rho = f(\xi^*_0, \cdots, \xi^*_n),$$

where f is a polynomial. Now $\omega \cdot (\eta^*)^\rho$ is an homogeneous element of degree ρ. Hence f is necessarily a form of degree ρ. It follows that $\omega = f/(\eta^*)^\rho$ is a polynomial in $\xi^*_0/\eta^*, \cdots, \xi^*_n/\eta^*$. Hence $\omega \subset \mathfrak{o}'$, and this proves our assertion.

Theorem 11′. *The manifold of singular points of a normal variety* V_r *is of dimension* $\leqq r - 2$.

This follows from Theorem 11 and from the fact that our V_r has just been proved to be normal in the affine space S'_n, *for any choice of the hyperplane at infinity.*

In this connection we wish to point out that *if, conversely, a* V_r *has the property that it is normal in the affine space* S'_n*, for every choice of the hyper-*

plane at infinity, then it does not yet necessarily follow that V_r is normal in the projective space P_n. We prove namely the following

THEOREM 13. *In order that a V_r in P_n be normal in the affine sense, for every choice of the hyperplane at infinity in P_n, it is necessary and sufficient that the conductor of the ring $\mathfrak{o}^* = K[\xi^*_0, \cdots, \xi^*_n]$ with respect to the integral closure $\bar{\mathfrak{o}}^*$ of $\mathfrak{o}^*$ (in the quotient field of $\mathfrak{o}^*$) divide a power of the 0-dimensional prime ideal $\mathfrak{p}^*_0 = (\xi^*_0, \cdots, \xi^*_n)$ (the vertex of the hypercone W_{r+1}).*

If $\mathfrak{o}^*$ is not itself integrally closed, the conductor must be then a primary ideal belonging to $\mathfrak{p}^*_0$.

Proof. The condition is sufficient. For let the conductor $\mathfrak{c}(\mathfrak{o}^*, \bar{\mathfrak{o}}^*)$ be a primary ideal $\mathfrak{q}^*$ belonging to $\mathfrak{p}^*_0$ and let σ be the exponent of $\mathfrak{q}^*$. If η^* is any linear form in $\xi^*_0, \cdots, \xi^*_n$ and if, as before, ω is any element in Σ which is integrally dependent on $\mathfrak{o}' = K[\xi^*_0/\eta^*, \cdots, \xi^*_n/\eta^*]$, then, for some integer ρ, the product $\omega \cdot (\eta^*)^\rho$ is in $\bar{\mathfrak{o}}^*$. Hence $\omega \cdot (\eta^*)^{\rho+\sigma} \subset \mathfrak{o}^*$, since $(\eta^*)^\sigma$ is in $\mathfrak{c}(\mathfrak{o}^*, \bar{\mathfrak{o}}^*)$. We conclude, as before, that $\omega \cdot (\eta^*)^{\rho+\sigma}$ is a form of degree $\rho + \sigma$ in $\xi^*_0, \cdots, \xi^*_n$ and that consequently $\omega \subset \mathfrak{o}'$, i. e. $\mathfrak{o}'$ is integrally closed in Σ.

The condition is necessary. Suppose that $\mathfrak{o}'$ is integrally closed in Σ. Let us consider any *homogeneous element* ω^* of $\bar{\mathfrak{o}}^*$. Let ω^* be homogeneous of degree ν, and let

$$(33) \qquad \omega^{*m} + a_1(\xi^*_0, \cdots, \xi^*_n)\omega^{*m-1} + \cdots + a_m(\xi^*_0, \cdots, \xi^*_n) = 0$$

be an equation of smallest degree for ω^* over $\mathfrak{o}^*$, with leading coefficient 1. The equation remains true if we apply any automorphism: $\xi^*_i \to t\xi^*_i$, $\omega^* \to t^\nu\omega^*$. The resulting equation must be an identity in t. Equating to zero the coefficient of $t^{\nu m}$, we get an equation similar to (33) but homogeneous of degree νm. Hence we may assume that $a_1, \cdots, a_m$ are homogeneous of degrees $\nu, 2\nu, \cdots, \nu m$ respectively, i. e. a_i is a form of degree $i\nu$ in $\xi^*_0, \cdots, \xi^*_n$.

From (33) it now follows that the element $\omega^*/(\eta^*)^\nu$, which is homogeneous of degree zero and is therefore an element of Σ, is integrally dependent on $\mathfrak{o}' = K[\xi^*_0/\eta^*, \cdots, \xi^*_n/\eta^*]$. Hence $\omega^*/(\eta^*)^\nu \subset \mathfrak{o}'$, i. e.

$$\omega^*/(\eta^*)^\nu = f(\xi^*_0/\eta^*, \cdots, \xi^*_n/\eta^*), \ f \text{ a polynomial.}$$

Clearing the denominator, we find

$$\omega^* \cdot (\eta^*)^\rho = g(\xi^*_0, \cdots, \xi^*_n), \qquad \rho \geqq 0,$$

where g is a form of degree $\rho + \nu$. Thus, given any linear form η^* in $\xi^*_0, \cdots, \xi^*_n$, there exists an integer ρ (perhaps depending on η^*) such that $\omega^* \cdot (\eta^*)^\rho \subset \mathfrak{o}^*$. In particular, let $\omega^* \cdot (\xi^*_i)^{\rho_i} \subset \mathfrak{o}^*$ and let $\rho' = 1 + \sum_{i=0}^{n} (\rho_i - 1)$.

Then it is clear that $\omega^* \cdot (\eta^*)^{\rho'} \subset \mathfrak{o}^*$ for any linear form η^* in $\xi^*_0, \cdots, \xi^*_n$. Hence $\omega^* \cdot \mathfrak{p}^*_0{}^{\rho'} \subset \mathfrak{o}^*$.

Now it is not difficult to see that every element in $\bar{\mathfrak{o}}^*$ is a sum of homogeneous elements of $\bar{\mathfrak{o}}^*$.[24] Since $\bar{\mathfrak{o}}^*$ possesses a finite $\mathfrak{o}^*$-basis, it follows that there also exists an $\mathfrak{o}^*$-basis in $\bar{\mathfrak{o}}^*$ consisting of homogeneous elements, say $\omega^*_1, \cdots, \omega^*_q$. We have just proved that for each ω^*_i we can find an integer σ_i such that $\omega^*_i \cdot \mathfrak{p}^*_0{}^{\sigma_i} \subset \mathfrak{o}^*$. If we put $\sigma = \max(\sigma_1, \cdots, \sigma_q)$, then we will have $\bar{\mathfrak{o}}^* \cdot \mathfrak{p}^*_0{}^{\sigma} \subset \mathfrak{o}^*$. Hence $\mathfrak{c}(\mathfrak{o}^*, \bar{\mathfrak{o}}^*) \supset \mathfrak{p}^*_0{}^{\sigma}$, and this proves our theorem.

We shall have occasion to point out later (see footnote 26) that there actually exist varieties which satisfy the condition of the last theorem and which yet are not normal.

19. There is an important connection between the notion of a normal variety V_r and the concept of a complete linear system of V_{r-1}'s on V_r. We begin with some simple remarks concerning *the order of* V_r.

Let $\xi^*_0, \xi^*_1, \cdots, \xi^*_n$ be homogeneous coördinates of the general point of V_r and let $\xi_i = \xi^*_i/\xi^*_0$ be the non-homogeneous coördinates. The order ν of V_r is ordinarily defined as the number of distinct intersections of V_r with a *general* $(n - r)$-dimensional subspace of P_n. Let

$$\eta_i = u_{i0} + u_{i1}\xi_1 + \cdots + u_{in}\xi_n, \qquad (i = 1, 2, \cdots, r),$$

where the u_{ij}'s are indeterminates. Clearly, the above definition of the order ν of V_r is equivalent with the following: ν is the relative degree of $\Sigma(u_{ij})$ with respect to the field $\mathsf{K}(\eta_1, \cdots, \eta_r; u_{ij})$ (see [7], p. 82). Now, with respect to this last field, there always exists in $\Sigma(u_{ij})$ a primitive element of the form $\eta_{r+1} = u_{r+1,0} + u_{r+1,1}\xi_1 + \cdots + u_{r+1,n}\xi_n$, where the coefficients $u_{r+1,i}$ are "non-special" constants in K. It follows that if all the u_{ij}'s are indeterminates, the irreducible equation $G(\eta_1, \cdots, \eta_{r+1}; u_{ij})$ between $\eta_1, \cdots, \eta_r, \eta_{r+1}$ is of degree ν in η_{r+1}. For reasons of symmetry it also must be of degree ν in each of the variables $\eta_1, \cdots, \eta_r$. Finally, since it is permissible to operate on $\eta_1, \cdots, \eta_r, \eta_{r+1}$ by a non-singular linear transformation, it follows that G *must be of degree* ν *in all the arguments* $\eta_1, \cdots, \eta_{r+1}$.

This is so as long as the u_{ij}'s are indeterminates. Now we specialize $u_{ij} \to u_{ij}^0 \subset \mathsf{K}$, $\eta_i \to \eta_i^0$. If the polynomial $G(\eta_1^0, \cdots, \eta_{r+1}^0; u_{ij}^0)$ in $\mathsf{K}[\eta_1^0, \cdots, \eta_{r+1}^0]$ does not vanish identically, then we get an algebraic rela-

[24] Since $\mathfrak{o}^* = \mathsf{K}[\xi_0^*, \xi_1^*, \cdots, \xi_n^*] = \mathsf{K}[\xi_0^*, \xi_0^*\xi_1, \cdots, \xi_0^*\xi_n] \subset \Sigma[\xi_0^*]$ and since $\Sigma[\xi_0^*]$ is integrally closed in $\Sigma^* = \Sigma(\xi_0^*)$, it follows that any element in $\bar{\mathfrak{o}}^*$ is a polynomial in ξ_0^* with coefficients in Σ. Let now ω^* be any element in $\bar{\mathfrak{o}}^*$, $\omega^* = \alpha_0 + \alpha_1\xi_0^* + \cdots, \alpha_s\xi_0^{*s}$, $\alpha_i \subset \Sigma$, and let $t_0 t_1, \cdots, t_s$ be $s + 1$ distinct constants. The automorphisms $\tau_{t_i}\colon \xi_j^* \to t_i\xi_j^*$ leave $\mathfrak{o}^*$ invariant. Hence the $s + 1$ elements $\omega_i^* = \tau_{t_i}\omega^* = \alpha_0 + t_i\alpha_1\xi_0^* + \cdots t^s\alpha_s\xi_0^{*s}$ also belong to $\bar{\mathfrak{o}}^*$. From this it follows that the homogeneous components $\alpha_0, \alpha_1\xi_0^*, \cdots, \alpha_s\xi_0^{*s}$ of ω^* are in $\bar{\mathfrak{o}}^*$.

tion between $\eta_1^0, \cdots, \eta_{r+1}^0$ which is of degree $\leqq \nu$. In the contrary case we still get an equation of degree $\leqq \nu$ between $\eta_1^0, \cdots, \eta_{r+1}^0$, provided that we specialize the u_{ij} one at a time and divide, when possible, by factors $u_{ij} - u_{ij}^0$. Consequently, *any* $r+1$ *linear polynomials* $\eta_1^0, \cdots, \eta_{r+1}^0$ *in the* ξ_i *satisfy an algebraic relation of degree* $\leqq \nu$. In particular, if r *of the elements* η_j^0 *are algebraically independent, then the irreducible algebraic relation between* $\eta_1^0, \cdots, \eta_{r+1}^0$ *is of degree* $\leqq \nu$. It is equal to ν, when the coefficients u_{ij}^0 are not special.

Let V'_r be another algebraic variety, birationally equivalent to V_r and lying in a projective n'-space $P_{n'}$, $n' > n$. Let $\xi^{*\prime}_0, \cdots, \xi^{*\prime}_{n'}$ be the homogeneous coördinates of the general point of V'_r. Here the $\xi^{*\prime}_j$ as well as the ξ^*_i are elements of the field Σ^*, a simple transcendental extension of Σ. The following is self-evident: V_r *is a projection of* V'_r *if* $\xi^*_0, \cdots, \xi^*_n$ *are proportional to linear forms in* $\xi^{*\prime}_0, \cdots, \xi^{*\prime}_{n'}$ *with coefficients in* K. Assuming, as it is permissible, that the ξ^*_i are linearly independent, we may so choose the coördinate system in $P_{n'}$ that $\xi^*_0, \cdots, \xi^*_n$ be proportional to $\xi^{*\prime}_0, \cdots, \xi^{*\prime}_n$.

We now define: The system of hyperplane sections of a V_r in P_n is said to be *complete*, if V_r is not the projection of a V'_r *of the same order* as V_r and *belonging* to a space $P_{n'}$ of dimension n' greater than n.[25]

When we say that V'_r *belongs* to $P_{n'}$ we mean that V'_r does not lie in any subspace of $P_{n'}$. Similarly we suppose that V_r belongs to P_n.

We now prove the following

THEOREM 14. *The system of hyperplane sections of a normal* V_r *is complete.*[26]

---

[25] The usual procedure is to give directly the general definition of a complete linear system of V_{r-1}'s on V_r. The property on which our definition is based becomes then a consequence of the definition as applied to the special case when the system of V_{r-1}'s is the system of hyperplane sections. We reverse the procedure. Dealing with an arbitrary linear system $|V_{r-1}|$ of V_{r-1}'s on a V_r, we would first transform our V_r into a V'_r on which the system $|V_{r-1}|$ is cut out by the hyperplanes, and then our definition is applicable, *provided* V_r *and* V'_r *are birationally equivalent.* The case in which the correspondence between V_r and V'_r is $(\alpha, 1)$, $\alpha > 1$, arises when the system $|V_{r-1}|$ on V_r is composite with an involution of degree α. This case would then require a separate treatment.

[26] This explains our term "normal." In the algebro-geometric literature a variety is called normal if its system of hyperplane sections is complete. However, it should be pointed out that while a variety, normal in our (arithmetic) sense, is also normal in the above geometric sense, the converse is not true. For instance, a curve may be normal in the geometric sense and still have singularities (example: a plane quartic of genus 2). Such a curve cannot be normal in the arithmetic sense, in view of Theorem 11′, section 18.

On the other hand, a curve may be free from singularities in the projective space and not be normal in the geometric sense—for instance—a rational space quartic. Such

Proof. Assume that V_r, of order ν and belonging to P_n, is the projection of birationally equivalent V'_r of the same order ν, lying in a $P_{n'}$, $n' > n$. Let $\xi^*_0, \cdots, \xi^*_n$ and $\xi^{*\prime}_0, \cdots, \xi^{*\prime}_{n'}$ be the homogeneous coördinates of the general point of V_r and of V'_r respectively. Then $\xi^*_0, \cdots, \xi^*_n$ are proportional to $\xi^{*\prime}_0, \cdots, \xi^{*\prime}_n$. Let $\xi^*_i/\xi^{*\prime}_i = \rho^* \subset \Sigma^*$, $i = 0, 1, \cdots, n$. The $n' + 1$ elements $\xi^*_0, \xi^*_1, \cdots, \xi^*_n, \rho^*\xi^{*\prime}_{n+1}, \cdots, \rho^*\xi^{*\prime}_{n'}$ are proportional to the coördinates $\xi^{*\prime}_j$, and moreover generate the field Σ^*, since $\Sigma^* = K(\xi^*_0, \cdots, \xi^*_n)$. Hence these $n' + 1$ elements can also be taken as homogeneous coördinates of the general point of V'_r. We may therefore assume that $\xi^*_i = \xi^{*\prime}_i$, $i = 0, 1, \cdots, n$. We drop the primes in the remaining $n' - n$ coördinates, so that now the coördinates of the general point of V'_r are $\xi^*_0, \cdots, \xi^*_n, \xi^*_{n+1}, \cdots, \xi^*_{n'}$.

To prove our theorem we have to show that V'_r actually belongs to P_n, i. e. *that* $\xi^*_{n+1}, \cdots, \xi^*_{n'}$ *are linearly dependent on* $\xi^*_0, \cdots, \xi^*_n$. Let $\xi_i = \xi^*_i/\xi^*_0$. Subject to a preliminary linear transformation on $\xi_1, \cdots, \xi_n$ we may assume that $\xi_1, \cdots, \xi_r$ are algebraically independent and that the relative degree $[\Sigma : K(\xi_1, \cdots, \xi_r)]$ is equal to ν. Let us now consider one of the elements $\xi_{n+1}, \cdots, \xi_{n'}$, say the element ξ_{n+1}. We can find constants $c_1, \cdots, c_n, c_{n+1}$; $c_{n+1} \neq 0$, such that $\bar{\xi}_{n+1} = c_1\xi_1 + \cdots + c_{n+1}\xi_{n+1}$ is a primitive element of Σ with respect to the field $K(\xi_1, \cdots, \xi_r)$. Let then

$$A_0(\xi_1, \cdots, \xi_r)\bar{\xi}^{\nu}_{n+1} + A_1(\xi_1, \cdots, \xi_r)\bar{\xi}^{\nu-1}_{n+1} + \cdots + A_\nu(\xi_1, \cdots, \xi_r) = 0$$

be the irreducible equation for $\bar{\xi}_{n+1}$ over $K(\xi_1, \cdots, \xi_r)$. Since V'_r is also of order ν, the above equation must be of degree ν in all the arguments $\xi_1, \cdots, \xi_r, \bar{\xi}_{n+1}$. Hence A_0 is a constant, and therefore $\bar{\xi}_{n+1}$ is integrally dependent on $K[\xi_1, \cdots, \xi_r]$. Since V_r is normal, the ring $K[\xi_1, \cdots, \xi_n]$ is integrally closed, and consequently $\bar{\xi}_{n+1} \subset K[\xi_1, \cdots, \xi_n]$. Since

$$\bar{\xi}_{n+1} = c_1\xi_1 + \cdots c_n\xi_n + c_{n+1}\xi_{n+1}, \qquad c_{n+1} \neq 0,$$

it follows that ξ_{n+1} is in $K[\xi_1, \cdots, \xi_n]$. Passing to the homogeneous coördinates we conclude that there exists an integer h_0 such that

$$\xi^*_{n+1}(\xi^*_0)^{h_0} \subset K[\xi^*_0, \cdots, \xi^*_n].$$

In a similar fashion we show the existence of an integer h_i such that

$$\xi^*_{n+1} \cdot \xi^{*h_i}_i \subset K[\xi^*_0, \cdots, \xi^*_n], \qquad (i = 0, 1, \cdots, n).$$

If then $h = \max(h_0, h_1, \cdots, h_n)$, then

$$\xi^*_{n+1} \cdot \xi^{*h}_i \subset K[\xi^*_0, \cdots, \xi^*_n], \qquad (i = 0, 1, \cdots, n).$$

---

a curve is normal in the affine space, for every choice of the hyperplane at infinity (see footnote 22), but since it is not normal in the geometric sense, it is *a fortiori* not normal in the arithmetic sense.

It follows that if we put $\lambda = (h-1)(n+1)+1$, and if η^* is an arbitrary form in $\xi^*_0, \cdots, \xi^*_n$, of degree λ, then $\xi^*_{n+1}\eta^* \subset K[\xi^*_0, \cdots, \xi^*_n]$. Since $\xi^*_{n+1}\eta^*$ is homogeneous of degree $\lambda+1$ (with respect to the automorphism $\xi^*_i \to t\xi_i$), it follows that $\xi^*_{n+1}\eta^*$ is a form in $\xi^*_0, \cdots, \xi^*_n$, of degree $\lambda+1$. Let $\omega_1, \cdots, \omega_N$ be the various power products of $\xi^*_0, \cdots, \xi^*_n$, of degree λ. If we apply the above result to $\eta^* = \omega_i$, we find:

$$\xi^*_{n+1}\omega_i = h_{i1}\omega_1 + \cdots + h_{iN}\omega_N, \qquad (i = 1, 2, \cdots, N),$$

where the h_{ij} are linear forms in $\xi^*_0, \cdots, \xi^*_n$. Hence $| h_{ij} - \delta_{ij}\xi^*_{n+1} | = 0$, where $\delta_{ij} = 0$ or 1 according as $i \neq j$ or $i = j$. It follows that ξ^*_{n+1} is integrally dependent on $K[\xi^*_0, \cdots, \xi^*_n]$, whence $\xi^*_{n+1} \subset K[\xi^*_0, \cdots, \xi^*_n]$, since V_r is normal. Now ξ^*_{n+1} is homogeneous of degree 1, *whence* ξ^*_{n+1} *is necessarily a linear form in* $\xi^*_0, \cdots, \xi^*_n$. In a similar fashion it follows that $\xi^*_{n+2}, \cdots, \xi^*_{n'}$ are linear forms in $\xi^*_0, \cdots, \xi^*_n$, and this completes the proof of our theorem.

20. We now proceed to establish the *existence* of normal varieties in any given class of birationally equivalent varieties. Specifically we shall show that for any given V_r in P_n it is possible to define a class of *derived normal varieties*—an extension of the analogous notion in affine spaces (see section 16).

Let V_r be an irreducible variety in P_n, and let $\Sigma = K(\xi_1, \cdots, \xi_n)$, $\Sigma^* = K(\xi^*_0, \xi^*_1, \cdots, \xi^*_n)$, where the ξ^*_i are the homogeneous coördinates of the general point of V_r, and $\xi_i = \xi^*_i/\xi^*_0$. Let $\mathfrak{o}^* = K[\xi^*_0, \cdots, \xi^*_n]$ and let $\bar{\mathfrak{o}}^*$ be the integral closure of $\mathfrak{o}^*$ in Σ^*. $\bar{\mathfrak{o}}^*$ is a finite integral domain, say

$$\bar{\mathfrak{o}}^* = K[\zeta^*_1, \cdots, \zeta^*_h].$$

Every element in $\bar{\mathfrak{o}}^*$ is a sum of homogeneous elements also belonging to $\bar{\mathfrak{o}}^*$, and an homogeneous element of $\bar{\mathfrak{o}}^*$ which is not a constant is necessarily of positive degree. Hence we may assume that each ζ^*_j is homogeneous and of degree $\delta_j > 0$. The integers $\delta_1, \cdots, \delta_h$ are not necessarily distinct; let $d_1, d_2, \cdots, d_q$ be the distinct integers among them and let d be their l. c. m. We consider in $\bar{\mathfrak{o}}^*$ the homogeneous elements whose degree is a given multiple of d, say σd. Every such element is a sum of power products of the ζ^*_j, each power product being an homogeneous element of $\bar{\mathfrak{o}}^*$, of degree σd. If in a given power product of the ζ^*_j there are g_1 factors of degree of homogeneity d_1, g_2 factors—of degree of homogeneity d_2, etc., then we must have

$$(34) \qquad g_1d_1 + g_2d_2 + \cdots + g_qd_q = \sigma d.$$

Thus the determination of the homogeneous elements of $\bar{\mathfrak{o}}^*$ of degree σd depends on finding all the non-negative solutions $g_1, \cdots, g_q$ of the above diophantine equation. Now suppose that $\sigma \geqq q$. Then, for at least one of the integers $g_1, \cdots, g_q$ we must have $g_i \geqq d/d_i$. If, say, $g_1 \geqq d/d_1$, then $(g_1 - d/d_1, g_2, \cdots, g_q)$ is a solution in non-negative integers of the equation

(34) with σ replaced by $\sigma - 1$, and the system $(g_1, g_2, \cdots, g_q)$ is the sum of the two systems

$$(g_1 - d/d_1, g_2, \cdots, g_q), \qquad (d/d_1, 0, \cdots, 0),$$

of which the second is a non-negative solution of the equation (34) for $\sigma = 1$. By repeated application of this reduction we conclude that if $\sigma \geqq q$, then every solution of (34) in non-negative integers g_j can be expressed as the sum of a non-negative solution of the equation

$$g_1 d_1 + \cdots + g_q d_q = (q-1)d$$

and of $\sigma - q + 1$ non-negative solutions of the equation $g_1 d_1 + \cdots g_q d_q = d$. If we now put $\delta = m \cdot (q-1)d$, m—an integer $\geqq 1$, we can assert that *every non-negative solution of the diophantine equation*

$$g_1 d_1 + \cdots + g_q d_q = \rho\delta, \quad \rho\text{—an arbitrary positive integer,}$$

is the sum of ρ *non-negative solutions of the diophantine equation*

$$g_1 d_1 + \cdots + g_q d_q = \delta.$$

This property is shared by any integer δ which is a multiple of $(q-1)d$. Actually it is very likely that d itself enjoys this property, but we have no proof for this conjecture.

Now consider any integer δ with the above property. Let $\omega^*_0, \omega^*_1, \cdots, \omega^*_m$, denote all the possible power products of $\zeta^*_1, \cdots, \zeta^*_h$, whose degree of homogeneity in $\bar{\mathfrak{o}}^*$ is equal to δ. By our choice of δ it follows that any power product of $\zeta^*_1, \cdots, \zeta^*_h$ whose degree of homogeneity in $\bar{\mathfrak{o}}^*$ is a multiple $\rho\delta$ of δ is necessarily a power product of $\omega^*_0, \cdots, \omega^*_m$, of degree ρ. Hence, *every element in* $\bar{\mathfrak{o}}^*$, *homogeneous of degree* $\rho\delta$, ρ*—an arbitrary positive integer, can be expressed as a form of degree* ρ *in* $\omega^*_0, \cdots, \omega^*_m$.

We shall call *character of homogeneity* of our V_r in P_n any integer δ which enjoys this last mentioned property. Any multiple of $(q-1)d$ is certainly a character of homogeneity of V_r.

Let δ be a character of homogeneity of V_r and let, as before, $\omega^*_0, \omega^*_1, \cdots, \omega^*_m$ be all the possible power products of $\zeta^*_1, \cdots, \zeta^*_h$, which are homogeneous elements of $\bar{\mathfrak{o}}^*$, of degree δ. The elements ω^*_i can be regarded as the homogeneous coördinates of the general point of an algebraic irreducible variety V'_r in a projective space $P_m(y'_0, y'_1, \cdots, y'_m)$. *The variety* V'_r *is birationally equivalent to* V_r. Namely, the quotients ω^*_i/ω^*_0, $i = 1, 2, \cdots, m$, are homogeneous elements of Σ^*, of degree zero, and hence are elements of Σ. Consequently $K(\omega^*_1/\omega^*_0, \cdots, \omega^*_m/\omega^*_0) \subseteqq \Sigma$. On the other hand, we have

$$\xi_i = \xi^*_i/\xi^*_0 = \xi^*_i \xi^{*\,\delta-1}_0 / \xi^{*\,\delta}_0.$$

The elements $\xi^*_i \xi^{*\,\delta-1}_0$ and $\xi^{*\,\delta}_0$ are homogeneous of degree δ, and consequently can be expressed as linear forms in $\omega^*_0, \omega^*_1, \cdots, \omega^*_m$. It follows that

$$(35) \qquad \xi_i = \frac{c_0^{(i)} + c_1^{(i)}\omega^*_1/\omega^*_0 + \cdots + c_m^{(i)}\omega^*_m/\omega^*_0}{c_0^{(0)} + c_1^{(0)}\omega^*_1/\omega^*_0 + \cdots + c_m^{(0)}\omega^*_m/\omega^*_0},$$

and consequently $\Sigma \subseteqq \mathsf{K}(\omega^*_1/\omega^*_0, \cdots, \omega^*_m/\omega^*_0)$. Hence the two fields $\mathsf{K}(\xi_1, \cdots, \xi_n)$ and $\mathsf{K}(\omega^*_1/\omega^*_0, \cdots, \omega^*_m/\omega^*_0)$ coincide, and this proves that V_r and V'_r are birationally equivalent. Note that the equations of the birational transformation between V_r in $P_n(y_0, \cdots, y_n)$ and V'_r in $P_m(y'_0, \cdots, y'_m)$ are, by (35), of the form

$$(36) \qquad \rho y_i = c_0^{(i)}y'_0 + c_1^{(i)}y'_1 + \cdots + c_m^{(i)}y'_m, \qquad (i = 0, 1, \cdots, n),$$

ρ—a factor of proportionality.

The linearity of these equations signifies that *V_r is a projection of V'_r*

We assert that the variety V'_r is normal (in its ambient projective space P_m). We have to show that the ring $\mathsf{K}[\omega^*_0, \cdots, \omega^*_m]$ is integrally closed in its quotient field. We first point out that since $\omega^*_0, \cdots, \omega^*_m$ are homogeneous elements of degree δ, every element ω^* in Σ^*, which depends integrally on $\omega^*_0, \cdots, \omega^*_m$, is a sum of homogeneous elements whose degrees are multiples of δ and which are also integrally dependent on $\omega^*_0, \cdots, \omega^*_m$. In view of the transitivity of integral dependence, the homogeneous components of ω^* are in $\bar{\mathfrak{o}}^*$. Hence they are forms in $\omega^*_0, \cdots, \omega^*_m$, and consequently $\omega^* \subset \mathsf{K}[\omega^*_0, \cdots, \omega^*_m]$, q. e. d.

21. In the construction of the normal variety V'_r there occur arbitrary elements, for instance the character of homogeneity δ. We thus get a whole class of normal varieties associated with V_r. Any variety of this class shall be called a *derived normal variety of V_r*. We wish to investigate the relationship between any two derived normal varieties of V_r.

A first arbitrary element in our construction is the choice of the elements $\zeta^*_1, \cdots, \zeta^*_h$, such that $\bar{\mathfrak{o}}^* = \mathsf{K}[\zeta^*_1, \cdots, \zeta^*_h]$. This choice affects the elements $\omega^*_0, \cdots, \omega^*_m$, which are the various power products of the ζ^*_j, of degree δ. However, since the elements ω^*_i always form a linear base for the homogeneous elements of degree δ in $\bar{\mathfrak{o}}^*$, it is clear that *two derived normal varieties of V_r belonging to one and the same character of homogeneity of V_r are projectively equivalent.*

Let now V'_r and V''_r be two derived normal varieties of V_r belonging to two distinct characters of homogeneity δ' and δ'' respectively. Let P_μ and P_ν be their ambient projective spaces respectively. Finally, let $\omega'_0, \cdots, \omega'_\mu$ (or $\omega''_0, \cdots, \omega''_\nu$) be the homogeneous coördinates of the general point of V'_r (or V''_r). We observe that if δ is a character of homogeneity of V_r, then any multiple of δ is a character of homogeneity. We may therefore consider the derived normal variety M_r of V_r belonging to the character of homogeneity $\delta'\delta''$. Let $\omega^*_0, \cdots, \omega^*_m$ be the homogeneous coördinates of M_r, its ambient

space being a P_m. The ω^* are forms of degree δ'' in $\omega'_0, \cdots, \omega'_\mu$, and every form of degree δ'' in $\omega'_0, \cdots, \omega'_\mu$ is necessarily a linear form in $\omega^*_0, \cdots, \omega^*_m$. It follows that M_r *is obtained from* V'_r *by referring projectively the hypersurfaces of order* δ'' *in* P_μ *to the hyperplanes in* P_m. (We assume that $\omega'_0, \cdots, \omega'_\mu$ are linearly independent, so that V'_r does not lie in a subspace of P_μ; we make similar hypotheses for V''_r in P_ν and M_r in P_m). In a similar manner it follows that M_r is obtained from V''_r by referring the hypersurfaces of P_ν of degree δ' to the hyperplanes of P_m. We have therefore the following

THEOREM 15. *The birational correspondence between* V'_r *and* V''_r *has the property that the linear system of sections of* V'_r *with the hypersurfaces of order* δ'' *of its ambient space is transformed into the linear system cut out on* V''_r *by the hypersurfaces of order* δ' *of its ambient space* P_ν.

This result implies, in particular, that the correspondence between V'_r and V''_r is $(1,1)$ *without exceptions.* It is free from fundamental elements on either variety. We connect up this result with the notion *of the quotient ring at a point of a* V_r.

Let $\xi^*_0, \cdots, \xi^*_n$ be the homogeneous coördinates of the general point of a V_r in P_n, and let $A(a_0, a_1, \cdots, a_n)$ be a given point of V_r. We define as the *quotient ring of the point* A, in symbols: $Q(A)$, the set of all elements $f(\xi^*)/g(\xi^*)$ in Σ (f, g—forms of like degree) such that $g(a) \neq 0$. In other words, $Q(A)$ consists of all elements of Σ which have a definite finite value at the point A. The quotient ring $Q(A)$ is independent of the choice of coördinates in P_n. In particular, if, say, $a_0 \neq 0$, and if we pass to non-homogeneous coördinates $\xi_i = \xi^*_i/\xi^*_0$, $x_i^0 = a_i/a_0$, then $Q(A)$ is the quotient ring of the prime ideal $\mathfrak{p}_0 = (\xi_1 - x_1^0, \cdots, \xi_n - x_n^0)$ in the ring $\mathfrak{o} = K[\xi_1, \cdots, \xi_n]$.

Let now W_r be another variety in a P_m, birationally equivalent to V_r, and let $\eta^*_0, \cdots, \eta^*_m$ be the homogeneous coördinates of the general point of W_r. Let

$$(37) \qquad \begin{cases} \rho\eta^*_i = f_i(\xi^*_0, \cdots, \xi^*_n), & (i = 0, 1, \cdots, m), \\ \sigma\xi^*_j = \phi_j(\eta^*_0, \cdots, \eta^*_m), & (j = 0, 1, \cdots, n), \end{cases}$$

be the equations of the birational transformation between V_r and W_r (the f_i-forms of like degree; similarly for the ϕ_j). Assume that the point A on V_r is not fundamental for the transformation. Then the quantities $f_i(a_0, a_1, \cdots, a_n)$ are not all zero, and there corresponds to A a unique point $B(b_0, b_1, \cdots, b_m)$ on W_r, where $\rho b_i = f_i(a_0, a_1, \cdots, a_n)$. If $g(\eta^*_0, \cdots, \eta^*_m)$ is a form such that $g(b_0, \cdots, b_m) \neq 0$, then it is clear that $g(f_0, \cdots, f_m)$ will be a form in $\xi^*_0, \cdots, \xi^*_n$ which is $\neq 0$ at A. It follows that $Q(B) \subseteqq Q(A)$. If we also assume that B is not a fundamental point on W_r, then we may conclude likewise that $Q(A) \subseteqq Q(B)$, whence $Q(A) = Q(B)$.

Finally, we point out that if $\mathfrak{o} = K[\xi_1, \cdots, \xi_n]$ is integrally closed in

its quotient field Σ, and if $\mathfrak{p}$ is any prime ideal in $\mathfrak{o}$, then the quotient ring $\mathfrak{o}_{\mathfrak{p}}$ is also intgrally closed in Σ. Reassuming, we may state the following

THEOREM 16. *The quotient ring of any point P of a normal V_r is integrally closed in the field of rational functions on V_r. The birational correspondence between the points of two derived normal varieties V'_r and V''_r of one and the same V_r, is $(1,1)$ without exceptions. The quotient rings of any two corresponding points P', P'' of V'_r and V''_r respectively, coincide.*

We conclude with a final important remark which clears up geometrically the relationship between a V_r and its derived normal varieties V'_r.

Let V'_r belong to the character of homogeneity δ and let $\omega^*_0, \cdots, \omega^*_m$ be the homogeneous coördinates of the general point of V'_r. Any form in $\xi^*_0, \cdots, \xi^*_n$ of degree δ, is necessarily a linear form in $\omega^*_0, \cdots, \omega^*_m$. Hence in the birational correspondence between V_r and V'_r, to the sections of V_r with the hypersurfaces of order δ, there correspond hyperplane sections of V'_r. In other words: if we denote by V_{r-1} the hyperplane sections of V_r, then the system of hyperplane sections of V'_r is the complete system $|\delta V_{r-1}|$. Thus we have the following

THEOREM 17. *The derived normal varieties V'_r of a given V_r are those on which the hyperplanes cut out the complete system $|\delta V_{r-1}|$, where δ is a character of homogeneity of V_r.*

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES

1. Grell, H., "Zur Theorie der Ordnungen in algebraischen Zahl- und Funktionenkörpen," *Mathematische Annalen*, vol. 97 (1927).
2. Krull, W., "Idealtheorie," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, IV, 3, Berlin, Springer (1935).
3. Muhly, H. T. and Zariski, O., "The resolution of singularities of an algebraic curve," *American Journal of Mathematics*, vol. 61, no. 1 (January, 1939).
4. Ore, O., "Uber den Zusammenhang zwischen den definierenden Gleichungen und der Idealtheorie in algebraischen Körpen, I," *Mathematische Annalen*, vol. 96 (1926).
5. Rückert, W., "Zum Eliminationsproblem der Potenzreihenideale," *Mathematische Annalen*, vol. 107 (1933).
6. Schmeidler, W., "Grundlagen einer Theorie der algebraischen Funktionen mehrerer Veränderlichen," *Mathematische Zeitschrift*, vol. 28 (1928).
7. van der Waerden, *Moderne Algebra*, II.
8. van der Waerden, "Zur algebraischen Geometrie III. Ueber irreduzibile algebraische Mannigfaltigkeiten," *Mathematische Annalen*, vol. 108 (1933).
9. Zariski, O.," Generalized weight properties of the resultant of $n+1$ polynomials in n indeterminates," *Transactions of the American Mathematical Society*," vol. 41, no. 2 (March, 1937).
10. Zariski, O., "Some results in the arithmetic theory of algebraic functions of several variables," *Proceedings of the National Academy of Sciences*, vol. 23, no. 7 (July, 1937).

ON SIMULTANEOUS EXPANSIONS OF ANALYTIC FUNCTIONS IN COMPOSITE POWER SERIES.*

By A. C. Burdette.

Introduction. The purpose of the first part of this paper is to prove that the system (1.1) of linear non-homogeneous functional equations (generalized difference equations) has solutions $g_k(x)$, $k=1,2,\cdots,n$, analytic in a certain finite region provided that the known functions $f_\nu(x)$, $\nu=1,2,\cdots,n$, are analytic in an appropriate region and provided that the system is non-singular in the sense defined in § 1.2.

In the second part of the paper we obtain facts concerning the convergence of series of the form (2.3) which Carmichael[1] has called composite power series and apply the local solution of the first part to obtain the simultaneous expansion of n analytic functions in composite power series.

The method employed is that used by Carmichael (C) in treating the foregoing problem in the case where the independent functions are integral. The results of the paper will include, as a special case, part of the results obtained by Carmichael in the paper cited.

I. On the Local Solution of a System of Linear Generalized Difference Equations with Constant Coefficients.

1.1. Formulation of the problem. We consider the problem of solving the system

$$\sum_{j=1}^{n} c_{\nu j} g_j(x-a_{\nu j}) = f_\nu(x), \qquad (\nu=1,2,\cdots,n), \tag{1.1}$$

of generalized difference equations where the functions $f_\nu(x)$, $\nu=1,2,\cdots,n$, are analytic in the neighborhood of some appropriate point, say $x=b$, and have there the power series expansions

$$f_\nu(x) = \sum_{j=0}^{\infty} \frac{s_{\nu j}}{j\,!}(x-b)^j, \qquad (\nu=1,2,\cdots,n). \tag{1.2}$$

For convenience in later work we define μ to be the least number satisfying the relations

* Received April 18, 1938.

[1] R. D. Carmichael, "Systems of linear difference equations and expansions in series of exponential functions," *Transactions of the American Mathematical Society*, vol 35 (1933), pp. 1-28. This paper will be denoted hereafter by (C).

$$\limsup_{j\to\infty}\left|\frac{s_{\nu j}}{j!}\right|^{1/j}\leq\mu,\qquad 0\leq\mu<\infty,\qquad(\nu=1,2,\cdots,n). \tag{1.3}$$

The coefficients $c_{\nu j}$ and the additive terms $a_{\nu j}$ are given complex constants which will later be subjected to certain negative conditions in order to avoid exceptional cases in the theory of the system. We seek to determine functions $g_k(x)$, $k=1,2,\cdots,n$, analytic in the neighborhood of $x=b$, which are solutions of the system (1.1).

Numerous writers [2] have treated related problems and special cases of the one set here. A recent paper by I. M. Sheffer [3] deals with a problem very closely related to the one under consideration. He obtains a local solution for a single equation of the type involved in the system (1.1). He also obtains an expansion for a single function similar to the simultaneous expansions of the second part of the present paper.

1.2. Symbolic Operators.[4] We define the operator $E(a)$ by the relations

$$E(a)\cdot f(x)=f(x+a),$$
$$\{\sum_{k=1}^{n}\alpha_k E(a_k)\}\cdot f(x)=\sum_{k=1}^{n}\alpha_k f(x+a_k),$$
$$\alpha E(a)\cdot\beta E(b)=\alpha\beta E(a+b),$$

where a, a_k, α, α_k, b, β are constants. These definitions give unique meaning to any polynomial combination of operators $E(a_k)$. Such a polynomial in operators E may be written as a linear function of suitably defined operators E. In particular, one may define such an operator by means of the symbolic determinant

$$\Delta\equiv|\,c_{\nu j}E(-a_{\nu j})\,|, \tag{1.4}$$

this being, by definition, the symbolic operator obtained by expanding the determinant formally as if its elements were ordinary algebraic quantities.

Any polynomial combination of operators E will be said to have the value zero when and only when the result of operating with it upon the function e^{xt} is identically zero when considered as a function of t.

Making use of the symbolic operators just introduced we may write the system (1.1) in the form

---

[2] See Carmichael (C), p. 2, for a list of some of the important papers in this connection.

[3] I. M. Sheffer, "A local solution of the difference equation $\Delta y(x)=F(x)$ and of related equations," *Transactions of the American Mathematical Society*, vol. 39 (1936), pp. 345-379.

[4] These symbolic operators are those used by Carmichael (C). We define them here for the convenience of the reader.

$$\sum_{j=1}^{n} c_{\nu j} E(-a_{\nu j}) g_j(x) = f_\nu(x), \qquad (\nu = 1, 2, \cdots, n). \tag{1.5}$$

The symbolic determinant in (1.4) will be called the symbolic determinant of the system (1.1), or (1.5). This determinant will be called singular when it has the value zero, otherwise it will be called non-singular.

We shall treat the system (1.1), or (1.5), only in the case when its determinant is non-singular. In that case we shall say the system is non-singular.

Since Δ is assumed to be non-singular, it may be written in the form

$$\Delta = \sum_{k=1}^{\sigma} c_k E(a_k), \tag{1.6}$$

where the constants c_k are different from zero and the constants a_k are different. The case $\sigma = 1$ is trivial; therefore in what follows we assume $\sigma > 1$.

1.3. Lemmas and definitions. The following lemma, used by Sheffer (l. c.), will prove useful in the work to follow:

LEMMA 1.1. *Let $\{\omega_k\}$ be a finite set of points in the finite plane. There exists a unique circle of smallest radius which covers the set, i. e., includes all points of the set in its interior or on its boundary.*

This unique circle associated with the set $\{\omega_k\}$ will be denoted by the symbol $\gamma\{\omega_k\}$.

With a given set $\{a_{\nu k}\}$ we shall associate the set $\{{}_0a_{\nu k}\}$ defined by ${}_0a_{\nu k} = a_{\nu k} - b$. If the center of the circle $\gamma\{a_{\nu k}\}$ is at the point b, the center of $\gamma\{{}_0a_{\nu k}\}$ is at zero; the radii of the two circles are the same. In the definitions and lemmas to follow we shall use the set $\{{}_0a_{\nu k}\}$.

We define the function $h(t)$ by the relation

$$h(t) = e^{-xt}\Delta e^{xt} = c_1 e^{{}_0a_1 t} + c_2 e^{{}_0a_2 t} + \cdots + c_\sigma e^{{}_0a_\sigma t}, \tag{1.7}$$

where Δ is the symbolic determinant of (1.5) for the set $\{{}_0a_{\nu k}\}$. If we factor $e^{{}_0a_1 t}$ out of $h(t)$ we have left

$$\bar{h}(t) = c_1 + c_2 e^{{}_0\bar{a}_2 t} + \cdots + c_\sigma e^{{}_0\bar{a}_\sigma t}, \tag{1.8}$$

where $c_1 \neq 0$ and ${}_0\bar{a}_k = {}_0a_k - {}_0a_1$. For a function of the type $\bar{h}(t)$ Carmichael (C) has proved the following lemma:

LEMMA 1.2. *There exists an infinite set of contours Γ_ν, $\nu = 0, 1, 2, \cdots$, such that there exists a positive number ϵ such that $|\bar{h}(t)| > \epsilon$ for all t on*

Γ_ν, $\nu = 0, 1, 2, \cdots$. *Moreover the contours* Γ_ν *are such that they all contain* $t = 0$ *in their interior; for* ν *greater than some preassigned number* N *the distance from* $t = 0$ *to a point* t *on* Γ_ν *is not less than* ν *nor greater than* $\nu + \beta$ *where* β *is a sufficiently large positive number; the length of* Γ_ν *bears a bounded ratio to* $2\pi\nu$.

We shall also have occasion to make use of the functions

$$e^{-{}_0a_1t}e^{-xt}A_{\nu k}e^{xt} = \sum_{i=1}^{P\nu k} d_i^{(\nu k)}e^{({}_0\alpha_i^{(\nu k)} - {}_0a_1)t},$$

where $A_{\nu k}$ is the cofactor of the element in the ν-th row and k-th column of the symbolic determinant Δ used in (1.7); $P\nu_k \leq (n-1)!$; and the sets of constants $d_i^{(\nu k)}$, ${}_0\alpha_i^{(\nu k)}$ are functions of the sets of constants $\{c_{\nu k}\}$, $\{a_{\nu k}\}$ respectively. Let α be defined as the least number satisfying the relation

$$(1.9) \qquad \alpha + 1 \geq |\, {}_0\alpha_i^{(\nu k)} - {}_0a_1 \,|, \qquad \text{(all } i, \nu, k).$$

Then by means of Lemma 1.2 we may state:

LEMMA 1.3. *For* t *on* Γ_ν, $\nu > N$,

$$|\, e^{-{}_0a_1t}e^{-xt}A_{mk}e^{xt} \,| < Me^{(\alpha+1)(\nu+\beta)}$$

for all m *and* k, *where* α *is defined by* (1.9) *and* β *is the constant of Lemma* 1.2.

1.4. Solution of the system (1.1). We shall consider (1.1) in the symbolic form (1.5). Let ρ and b denote, respectively, the radius and center of $\gamma\{a_{\nu k}\}$. We may without loss of generality assume $b = 0$. This amounts to considering the system (1.5) formed on the set $\{{}_0a_{\nu k}\}$. However, for convenience, we retain the same notation.

Form the functions

$$(1.10) \quad g_k(x) = \frac{1}{2\pi i}\sum_{j=0}^{\infty}\sum_{\nu=1}^{n}\int_{\Gamma_j} s_{\nu j}(e^{-xt}A_{\nu k}e^{xt})\,\frac{e^{xt}dt}{h(t)\,t^{j+1}}, \quad (k = 1, 2, \cdots, n),$$

where the coefficients $s_{\nu j}$ are those appearing in (1.2); $h(t)$ is the function defined in (1.7); the contours Γ_j are those of Lemma 1.2; $A_{\nu k}$ is the cofactor of the element in the ν-th row and k-th column of Δ; and $i = \sqrt{-1}$. By application of the Lemmas 1.2 and 1.3 one readily obtains that the absolute value of the j-th term of (1.10) is dominated by

$$Me^{(\alpha+1)(j+\beta)}e^{|x|(j+\beta)}j!\cdot j^{-j}\sum_{\nu=1}^{n}\frac{|\, s_{\nu j} \,|}{j!}$$

for $j > N$. Hence for μ sufficiently small (μ being the constant defined by (1.3)) the series (1.10) converge absolutely for

$$| x | < \log e^{-\alpha}\mu^{-1}$$

and therefore define functions analytic in the neighborhood of $x = 0$.

The series (1.10) are readily seen to afford a formal solution of the system (1.5). If we impose the further restriction on μ, i. e., on the series (1.2), that

$$\log e^{-\alpha}\mu^{-1} > \rho$$

then the circles of convergence of the series defining $g_k(x - a_{\nu k})$, $\nu, k = 1, 2, \cdots, n$, have a region in common in the neighborhood of $x = 0$. Therefore, under these conditions, the functions $g_k(x)$ defined by (1.10) afford an actual solution of the system (1.5), or (1.1), and we have:

THEOREM 1.1. *Let ρ and b denote, respectively, the radius and center of the circle $\gamma\{a_{\nu k}\}$ where the constants $a_{\nu k}$ are such that $b = 0$, and let α denote the constant defined in* (1.9). *If the functions $f_\nu(x)$, $\nu = 1, 2, \cdots, n$, are analytic throughout the interior of the circle $| x | = \mu^{-1}$ where μ is such that $\log e^{-\alpha}\mu^{-1} > \rho$, the functions $g_k(x)$, $k = 1, 2, \cdots, n$, defined by* (1.10) *are analytic in the neighborhood of $x = 0$ and satisfy the non-singular system* (1.1) *in a region containing that point in its interior. This region includes the open region common to the set of circles $| x - a_{\nu k} | = \log e^{-\alpha}\mu^{-1}$, $\nu, k = 1, 2, \cdots, n$.*[5]

As a consequence of this theorem we have:

THEOREM 1.2. *Let ρ and b denote, respectively, the radius and center of the circle $\gamma\{a_{\nu k}\}$, and let α denote the constant defined in* (1.9). *If the functions $f_\nu(x)$, $\nu = 1, 2, \cdots, n$, are analytic throughout the interior of the circle $| x - b | = \mu^{-1}$ where μ is such that $\log e^{-\alpha}\mu^{-1} > \rho$, there exists functions $g_k(x)$, $k = 1, 2, \cdots, n$, analytic in the neighborhood of $x = b$ which satisfy the non-singular system* (1.1) *in a region containing that point in its interior. This region includes the open region common to the set of circles*

$$| x - a_{\nu k} | = \log e^{-\alpha}\mu^{-1}, \qquad (\nu, k = 1, 2, \cdots, n).$$ [6, 7]

---

[5] It is evident that this theorem gives the best results when the a_1 used in defining α is a member of the set $\{a_k\}$ such that its distance from the center of the circle $\gamma\{{}_0a_i{}^{(\nu k)}\}$ is a minimum. This is purely a question of the choice of notation.

[6] If $\mu = 0$ the functions $f_\nu(x)$ are integral and we obtain Carmichael's (C) result p. 14.

[7] By the same method a similar theorem can be obtained for the more general system

$$\sum_{j=1}^{n} \{ \sum_{r=1}^{m_{\nu j}} C_{\nu j;r}\, g_j(x - a_{\nu j;r}) \} = f_\nu(x) \qquad (\nu = 1, 2, \cdots n).$$

The only changes necessary in such a development are obvious modifications of the definitions of the auxiliary functions and constants.

II. Simultaneous Expansions of Analytic Functions in Composite Power Series.

2.1. Formulation of the problem. For $n > 1$ we consider the problem of expanding n functions $f_\nu(x)$, $\nu = 1, 2, \cdots, n$, analytic in the neighborhood of a point $x = b$, simultaneously in composite power series, i. e., we consider the problem of representing these functions in the form

$$f_\nu(x) = \sum_{k=0}^{\infty} \sum_{j=1}^{n} \alpha_{kj} c_{\nu j} (x - a_{\nu j})^k, \qquad (\nu = 1, 2, \cdots, n), \tag{2.1}$$

where the coefficients α_{kj} are independent of both x and ν. We seek expansions (2.1) in which the coefficients α_{kj} are such that the series in the equations

$$g_j(x) = \sum_{k=0}^{\infty} \alpha_{kj} x^k, \qquad (j = 1, 2, \cdots, n), \tag{2.2}$$

converge in the neighborhood of $x = 0$. We subject the constants $c_{\nu j}$, $a_{\nu j}$ to the condition that the determinant $\Delta(t)$, whose element in the ν-th row and j-th column is $c_{\nu j} e^{-a_{\nu j} t}$, shall not be identically zero in t.

Under the conditions named we shall show that for suitable functions $f_\nu(x)$ such expansions exist.

2.2. Convergence of composite power series. Before continuing with the problem set in the preceding section we shall obtain certain convergence properties of series of the form

$$\sum_{\nu=0}^{\infty} \sum_{k=1}^{n} \alpha_{\nu k} c_k (x - a_k)^\nu. \tag{2.3}$$

We assume that the a_k are n different constants; that $c_k \neq 0$, $k = 1, 2, \cdots, n$; that an infinite number of each set $\alpha_{\nu k}$, $k = 1, 2, \cdots, n$, are different from zero; and that

$$\limsup_{\nu \to \infty} | \alpha_{\nu k} c_k |^{1/\nu} = r_k, \qquad (0 < r_k < \infty).$$

There is no loss in assuming $r_k > 0$ because if any $r_k = 0$ the corresponding series

$$\sum_{\nu=0}^{\infty} \alpha_{\nu k} c_k (x - a_k)^\nu$$

converges throughout the finite plane and consequently has no effect on the region of convergence of (2.1).

Let the series of circles

$$|x - a_k| = r_k^{-1}, \qquad (k = 1, 2, \cdots, n)$$

be called the fundamental set of circles, and let the set of points composing the boundaries of the set of circles defined by the relations

$$r_i\,|x - a_i| = r_j\,|x - a_j|, \qquad (i, j = 1, 2, \cdots, n),$$

be called the exceptional set.

By means of the convergence properties of ordinary power series and by considering the absolute value of the terms of (2.3) over a subsequence of $\nu = 0, 1, 2, \cdots$, we are led to the theorem:

THEOREM 2.1. *Let C denote the closed region consisting of all points common to all members of the fundamental set of circles (boundaries included). Then*

(*i*) *The series* (2.3) *converges absolutely at each interior point of C; it converges absolutely and uniformly in any closed region interior to C;*

(*ii*) *The series* (2.3) *diverges at every point exterior to C and not in the exceptional set;*

(*iii*) *If the region C is vacuous there are no points of convergence of the series* (2.3) *save possibly points of the exceptional set.*

Examples can readily be constructed to show that there may be points of convergence in the exceptional set and outside C. Examples can also be constructed to show that (iii) of the theorem may exist.

2.3. Expansions of analytic functions. If we employ the notation defined in (2.2) we may write (2.1) in the form

$$f_\nu(x) = \sum_{j=1}^{n} c_{\nu j} g_j(x - a_{\nu j}), \qquad (\nu = 1, 2, \cdots, n). \tag{2.4}$$

Suitable solutions of this system evidently lead through (2.2) to the required expansions (2.1). The condition put on $\Delta(t)$ in § 2.1 is just that required to make the results of the first part applicable to the system (2.4). Let the notation of § 1.1 be carried over to this section. Then we may state:

THEOREM 2.2. *Let $a_{\nu j}$, $c_{\nu j}$, $\nu, j = 1, 2, \cdots, n$, be two sets of complex constants such that $\Delta(t) \not\equiv 0$; let ρ and b denote, respectively, the radius and*

center of the circle $\gamma\{a_{\nu k}\}$; let α be the constant defined in (1.9). *If the functions $f_\nu(x)$, $\nu = 1, 2, \cdots, n$, are analytic throughout the interior of the circle $|x - b| = \mu^{-1}$ where μ is such that $\log e^{-\alpha}\mu^{-1} > \rho$, they have simultaneous expansions of the form* (2.1) *valid in a region including $x = b$ in its interior. This region includes the open region common to the set of circles*

$$|x - a_{\nu k}| = \log e^{-\alpha}\mu^{-1}, \qquad (\nu, k = 1, 2, \cdots, n).^{8}$$

The formula (1.10)[9] affords a means of obtaining suitable coefficients α_{kj} to be used in expansions (2.1). If we define $\Delta_{\nu k}(t)$ to be the cofactor of the element in the ν-th row and j-th column of $\Delta(t)$, then coefficients α_{kj} are given by

$$\alpha_{kj} = \frac{1}{2\pi i(k!)} \sum_{m=0}^{\infty} \int_{\Gamma_m} \Big(\sum_{\nu=1}^{n} s_{\nu m}\Delta_{\nu j}(t)\Big) \frac{dt}{\Delta(t)t^{m-k+1}},$$
$$(j = 1, 2, \cdots, n;\ k = 0, 1, 2, \cdots).$$

UNIVERSITY OF CALIFORNIA.

---

[8] By means of the generalization indicated in [7] it is clear that simultaneous expansions of the form

$$f_\nu(x) = \sum_{k=0}^{\infty} \sum_{j=1}^{n} \Big\{\sum_{r=1}^{m_{\nu j}} a_{kj\,;\,r}c_{\nu j}\,(x - a_{\nu j\,;\,r})^k\Big\}, \qquad v = 1, 2, \cdots, n,$$

may be obtained.

[9] It is to be recalled that (1.10) is formed with respect to the set $\{{}_0a_{\nu k}\}$.

SOME THEOREMS ON QUASI-ANALYTICITY FOR FUNCTIONS OF SEVERAL VARIABLES.*

By S. Bochner and A. E. Taylor.[1]

Introduction. The notion of quasi-analyticity may be described roughly as follows. A class of functions defined in a fixed domain is said to be quasi-analytic if within the class an individual function is completely determined by its behavior in an arbitrarily small sub-region of its domain of definition, or, what is sometimes the same thing, if the function is determined by the knowledge of its value and that of all its derivatives at a single point.

In the present paper we shall exhibit various ways of realizing, for definite classses of functions of several variables, this notion of quasi-analyticity. All of our results depend upon the following theorem of Denjoy and Carleman, which has to do with functions of one real variable.[2]

Theorem D. *Let f be a real function of the real variable x, defined and possessing derivatives of all orders on the interval $a \leqq x \leqq b$. Then if $f^{(n)}(a) = 0$ $(n = 0, 1, \cdots)$ and if for all values of x the inequalities*

$$|f^{(n)}(x)| \leqq m_n \qquad (n = 0, 1, \cdots)$$

are satisfied, where the constants m_n are such that the series

$$\Sigma\, m_n^{-1/n}$$

is divergent, then f is identically zero.

It is evident from this that for a fixed sequence of positive constants $\{m_n\}$ such that $\Sigma m_n^{-1/n}$ is divergent the class of functions f defined on (a, b) and such that

$$|f^{(n)}(x)| \leqq A^n m_n \qquad (n = 0, 1, \cdots)$$

where A is a constant depending only on f, is quasi-analytic in the sense that if f and g are in the class, and $f^{(n)}(a) = g^{(n)}(a)$, then $f \equiv g$.

When one seeks to establish some similar result for functions of several variables a number of different modes of procedure present themselves. If we

* Received July 28, 1938.

[1] National Research Fellow.

[2] T. Carleman, *Les Fonctions Quasi Analytiques*, Paris (1926), p. 20.

consider, to begin with, Euclidean spaces only, we may, with a fixed domain R, restrict the magnitude of the function and all its partial derivatives in R. Or we may perhaps require restrictions on certain combinations of these derivatives only, that is to say, on certain differential operators applicable to f in R—the successive iterates of the Laplacian, for example. Again, the conditions $f^{(n)}(a) = 0$ may have as their analogues in the generalization conditions on the function and its derivatives at a single point, or they may be replaced by similar conditions relating to a point set in R with the property that a function which is analytic in R and vanishes on the set must vanish identically. We have made use of each of these alternatives, and it is on the basis of them that the paper is divided into two parts, the second part dealing with certain differential operators, especially the Laplacian. The theorems of the second part, it will be seen, are essentially theorems on functions of several variables, being genuinely different in character from those of the first part.

We have also considered Riemannian spaces; in this case likewise we have resorted to generalizations of two different kinds. Theorems 3, 4, 5, 7 (Part I) and Theorem 10 (Part II) deal with Riemannian spaces. In particular, Theorem 10 is concerned with functions defined on a space of constant positive curvature, the unit sphere

$$x_1^2 + \cdot \cdot \cdot + x^2_{k+1} = 1$$

in Euclidean space of $k + 1$ dimensions, $k \geqq 2$. In its formal statement Theorem 10 differs from Theorem D merely by the substitution of $\Delta^n f$ for $f^{(n)}(x)$, Δ being the Laplace-Beltrami operator for the sphere, and the substitution of a point set of specified character for the single point $x = a$.

In order to avoid repetitions in the statement of our theorems, all of which have the same general form, we shall agree at the outset on the following conventions: a) f shall denote a real, continuous functions defined in some domain (= open set) of a space; it shall possess continuous derivatives of all orders with respect to the coördinates used. b) $\{m_n\}$ $n = 0, 1, \cdot \cdot \cdot$ shall denote a sequence of positive constants, arbitrary except for the conditions imposed in the statement of the theorem in which it occurs. c) Statements involving the index n shall hold for $n = 0, 1, \cdot \cdot \cdot$ unless the contrary is explicitly asserted. d) If $(x_1, \cdot \cdot \cdot, x_k)$ are the coördinates of a point, we shall frequently refer to it as the point x; if f is a function of $x_1, \cdot \cdot \cdot, x_k$ we shall often express this by the notation $f(x)$. e) E_k shall denote the Eucildean arithmetic space of k dimensions; V_k shall denote a Riemannian space of k dimensions. We shall require merely that the components of the fundamental metric tensor g_{ij} for the space V_k be of class C^∞; that is, they admit continuous partial derivatives of all orders. We stress the fact that the g's need not be analytic.

Part I.

1. If f is defined over a domain of E_k we introduce the symbols

$$(1.1)\qquad \begin{aligned} D_0(f,x) &= |f(x)| \\ D_n(f,x) &= \left\{ \sum \left| \frac{\partial^n f}{\partial x_{\alpha_1}\cdots\partial x_{\alpha_n}} \right|^2 \right\}^{1/2} \end{aligned} \qquad (n=1,2,\cdots)$$

where $\alpha_1,\cdots,\alpha_n$ are summed independently from 1 to k.

THEOREM 1. *Let f be defined in a connected domain R in E_k, and let x^0 be an interior point of R. Then the conditions*

a) $D_n(f,x) \leqq m_n$ $\quad x$ in R
b) $D_n(f,x^0) = 0$
c) $\sum m_n^{-1/n} = \infty$ $\quad$ (the series is divergent)

imply that f is zero throughout R.

Proof. Selecting an arbitrary straight line emanating from x^0 and a point x' on this line such that the segment x^0x' lies wholly in R, we have as a parametric representation of this segment

$$x_i(s) = a_i s + x_i^0 = \frac{(x'_i - x_i^0)}{d} s + x_i^0$$

where d is the distance from x^0 to x' and s is the distance from x^0 to $x(s)$, so that

$$(1.2)\qquad \sum_1^k a_i^2 = 1.$$

Define

$$F(s) = f(x_1(s),\cdots,x_k(s)) \equiv f(x(s)).$$

Then clearly

$$F^{(n)}(s) = \sum a_{\alpha_1}\cdots a_{\alpha_n} f_{\alpha_1\ldots\alpha_n}(x(s))$$

where

$$f_{\alpha_1\ldots\alpha_n}(x) = \frac{\partial^n f}{\partial x_{\alpha_1}\cdots\partial x_{\alpha_n}}.$$

By (1.2) and the Cauchy-Schwarz inequality we find

$$|F^{(n)}(s)| \leqq D_n(f,x(s)) \leqq m_n \qquad 0 \leqq s \leqq d$$

and so also $F^{(n)}(0) = 0$. Therefore by Theorem D, $F(s) = 0$ and in particular $f(x') = 0$.

Since R is connected it is obvious that a finite number of applications of

the above reasoning allows us to conclude, as in the process of analytic continuation, that f is zero at all points of R.

Theorem 1 is perhaps the most obvious generalization of Theorem D. The generalization may be pushed even further, to include functions defined on an arbitrary Banach space[3] E. For this purpose we introduce the following considerations. If E is a Banach space with elements $x, y, \cdots$, and f is a real function defined in a domain R in E, we shall say that f is differentiable arbitrarily often along all lines in R if the quantities

$$\delta^n f(x;y) = \left[\frac{\partial^n f(x+sy)}{\partial s^n}\right]_{s=0} \tag{1.3}$$

exist for each x in R and each y in E. It is readily seen that if E is the Euclidean space E_k we have

$$\delta^n f(x;y) = \sum \frac{\partial^n f}{\partial x_{a_1} \cdots \partial x_{a_n}} y_{a_1} \cdots y_{a_n}$$

and consequently

$$|\delta^n f(x;y)| \leqq D_n(f,x)\{\sum y_i^2\}^{n/2}.$$

This leads us at once to the statement of our next theorem.

THEOREM 2. *Let f be defined in a connected domain R of a Banach space E, and let f be differentiable arbitrarily often along all lines in R. Then if x^0 is a point in R the conditions*

$$\begin{array}{ll} \text{a)} & |\delta^n f(x;y)| \leqq m_n \|y\|^n \quad x \text{ in } R,\ y \text{ in } E \\ \text{b)} & \delta^n f(x^0;y) = 0 \\ \text{c)} & \sum m_n^{-1/n} = \infty \end{array}$$

imply that f is zero throughout R.

Proof. If x' is chosen as in the proof of Theorem 1, and we put

$$F(s) = f(x^0 + s(x' - x^0)) \qquad 0 \leqq s \leqq 1$$

then

$$F^{(n)}(s) = f(x^0 + s(x' - x^0); x' - x^0)$$

and by conditions a)–c) and Theorem D we conclude that $F(s) \equiv 0$. The rest of the proof is then obvious.

2. In order to formulate theorems of the desired type for Riemannian spaces it is first of all necessary to devise something corresponding to the

[3] For the definition of such spaces see Benach, *Opération Linéaires*, Warsaw (1932), Chapter V.

quantities $D_n(f, x)$ entering in Theorem 1. It is desirable to define $D_n(f, x)$ in invariant form; covariant differentiation naturally suggests itself. However, if we keep formulae (1. 1) in mind and think in terms of normal coördinates, it is another process which occurs most readily as the basis for our procedure. This process is that of *extension.* All that we need of it for our purposes is the following. If f is a function defined in V_k with x-coördinate system, and x_0 is a fixed point, we introduce normal coördinates $(y^1, \cdots, y^k)$ with origin at this point. Then if

$$f^*(y^1, \cdots, y^k) \equiv f(x^1, \cdots, x^k) \tag{2.1}$$

we define

$$f_{a_1 \ldots a_n}(x_0) = \left(\frac{\partial^n f^*}{\partial y^{a_1} \cdots \partial y^{a_n}}\right)_{y=0}. \tag{2.2}$$

The functions $f_{a_1 \ldots a_n}$ defined by (2. 2) constitute the components of an absolute tensor, covariant of rank n.[4]

Since we shall make extensive use of normal coördinates it will be convenient to assemble for reference the principal facts on which we shall rely. In the neighborhood of any given point of V_k it is possible to introduce normal coördinates y^i with origin at the point in question. If f is a scalar function defined over a portion of V_k we shall use f^* to denote this scalar as a function of normal coördinates about a chosen point. We shall use ψ_{ij} to denote the components, in normal coördinates, of the fundamental metric tensor of the space, and $C_{\alpha\beta}{}^{\gamma}$ to denote the Christoffel symbols in normal coördinates. The normal coördinates may be so altered by a linear transformation that at the origin of the system

$$\psi_{ij}(0) = \psi^{ij}(0) = \delta_j{}^i. \tag{2.3}$$

The equations of a geodesic through the origin are

$$y^i = a^i s \tag{2.4}$$

where s is the distance along the geodesic, measured from the origin, and the a's are constants satisfying the relation

$$\sum_1^k a^{i^2} = 1. \tag{2.5}$$

Along the geodesic the following fundamental relations hold.[5]

---

[4] T. Y. Thomas, *The Differential Invariants of Generalized Spaces,* Cambridge University Press (1934), p. 97.

[5] T. Y. Thomas, *op. cit.,* p. 85-87.

$$\text{(2.6)}\qquad \begin{cases} s^2 = \psi_{ij}(0)y^i y^j = \psi_{ij}(y)y^i y^j \\ \psi_{ij}(0)y^i = \psi_{ij}(y)y^i \end{cases}$$

$$\text{(2.7)}\qquad C_{\alpha\beta}{}^{\gamma} y^{\alpha} y^{\beta} = 0.$$

With the quantities (2.2) we define

$$\text{(2.8)}\qquad \begin{aligned} D_0(f,x) &= |f(x)| \\ D_n(f,x) &= \{g^{\alpha_1\beta_1}\cdots g^{\alpha_n\beta_n} f_{\alpha_1\ldots\alpha_n} f_{\beta_1\ldots\beta_n}\}^{1/2}. \end{aligned}$$

Using normal coördinates with origin at x, and taking account of (2.2), (2.3) we see that

$$D_n(f,x) = \left\{ \Sigma \left| \frac{\partial^n f^*}{\partial y^{\alpha_1}\cdots\partial y^{\alpha_n}} \right|^2_{y=0} \right\}^{1/2}.$$

Thus $D_n(f,x)$ is a real, non-negative scalar.

We are now in a position to enunciate a theorem.

THEOREM 3. *Let f be defined in a connected domain R of V_k. Then, x_0 being a point in R and $D_n(f,x)$ being given by* (2.8), *f is identically zero provided that conditions* a)–c) *of Theorem* 2 *are fulfilled.*

Proof. Consider a geodesic sphere of radius r in R with center at x_0, and a fixed geodesic passing through x_0. Along this geodesic f becomes a function $F(s)$ of the distance s from x_0. Let x_1 be an arbitrary point on the given geodesic, within the geodesic sphere. Then if we choose normal coördinates with origin at x_1 we have equations similar to (2.4) for the geodesic, s being replaced by $s - s_1$, where s_1 is the value of s at $x = x_1$. Clearly

$$F^{(n)}(s_1) = a^{\alpha_1}\cdots a^{\alpha_n}\left(\frac{\partial^n f^*}{\partial y^{\alpha_1}\cdots\partial y^{\alpha_n}}\right)_{y=0}$$

and therefore, by (2.5) and the Schwarz inequality

$$|F^{(n)}(s_1)| \leqq D_n(f,x_1) \leqq m_n.$$

Hence also $F^{(n)}(0) = 0$. The reasoning then follows that of the previous proofs.

THEOREM 4. *The hypothesis and conclusion here are the same as in Theorem 3, except that in the definition of $D_n(f,x)$* [formula (2.8)], *$f_{\alpha_1\ldots\alpha_n}$ shall signify the components of the n-th successive covariant derivative of f.*

Proof. Consider as before a geodesic sphere with center at x_0, and the function $F(s)$. It is clear that it is sufficient to prove that for an arbitrary geodesic through x_0 the inequalities

$$(2.9) \qquad | F^{(n)}(s) | \leqq D_n(f, x)$$

are satisfied, where x and s refer to the same point. We shall first turn our attention to a lemma independent of the above considerations.

LEMMA 1. *Let S be a region in V_k which is entirely covered by a system of normal coördinates y^i with origin in the region. Let f be a scalar defined in S. If the geodesics through the origin are defined by* (2.4) *subject to condition* (2.5), *so that the variables (s, a) furnish a fixed set of "polar" coördinates covering S, we define*

$$F(s; a) = f^*(y).$$

In the region S with the origin excluded the derivatives

$$F^{(n)}(s; a) \equiv \frac{\partial^n F(s; a)}{\partial s^n}$$

are then seen to be scalars. As such they are given by the formulae

$$(2.10) \qquad F^{(n)}(s; a) = \psi^{i_1 a_1} \cdots \psi^{i_n a_n} \Omega_{,i_1} \cdots \Omega_{,i_n} f^*_{a_1 \ldots a_n}$$

where $\Omega_{,i}$ are the components of the covariant derivative of the scalar

$$(2.11) \qquad \Omega = \{\psi_{ij}(y) y^i y^j\}^{1/2}$$

and $f^*_{a_1 \ldots a_n}$ are the components of the n-th successive covariant derivative of f^*.

Proof of the lemma. We have

$$F'(s; a) = a^a \frac{\partial f^*}{\partial y^a} = \frac{y^a}{s} \frac{\partial f^*}{\partial y^a}.$$

Now

$$\Omega_{,i} = \frac{\partial \Omega}{\partial y^i}, \qquad f^*_a = \frac{\partial f^*}{\partial y^a}$$

and, remembering relations (2.6)

$$\frac{\partial \Omega}{\partial y^i} = \frac{\psi_{ij}(y) y^j}{\Omega} = \frac{\psi_{ij}(0) y^j}{\Omega}$$

from which follows

$$(2.12) \qquad \psi^{ia} \frac{\partial \Omega}{\partial y^i} = \frac{\psi^{ia} \psi_{ij} y^j}{\Omega} = \frac{y^a}{s}.$$

Therefore (2.10) has been proved for $n = 1$.

Replacing the scalar $F(s; a)$ by $F'(s; a)$ and applying the result just obtained we get

$$F^{(2)}(s;a) = \psi^{i\alpha}\Omega_{,i}(\psi^{j\beta}\Omega_{,j}f^*{}_\beta)_{,\alpha}$$
$$= \psi^{i\alpha}\psi^{j\beta}(\Omega_{,i}\Omega_{,j}f^*{}_{\beta\alpha} + \Omega_{,i}\Omega_{,j,\alpha}f^*{}_\beta).$$

But this will agree with (2.10) if

$$(2.13) \qquad \psi^{i\alpha}\psi^{j\beta}\Omega_{,i}\Omega_{,j,\alpha} = 0 \qquad (\beta = 1, \cdots, k).$$

To prove (2.13) we write the left side in the form

$$\frac{y^\alpha}{s}\left(\frac{y^\beta}{s}\right)_{,\alpha} = \frac{y^\alpha}{s}\left\{\frac{\partial}{\partial y^\alpha}\left(\frac{y^\beta}{s}\right) + \frac{y^\gamma}{s}C_{\gamma\alpha}{}^\beta\right\}$$

which is possible because of (2.12) and the fact that the covariant derivative of the tensor ψ_{ij} vanishes. But the last expression above on the right is zero because of the fundamental relations (2.6), (2.7), as is easily verified.

Proceeding in this way, by repeated use of (2.10) for $n = 1$, and the relation (2.13), we conclude (2.10) in general.

If in (2.10) we fix the a's and allow s to approach zero we find, in view of (2.12) and (2.4), (2.5), that

$$(2.14) \quad \lim_{s\to 0} |F^{(n)}(s;a)| = |a^{\alpha_1}\cdots a^{\alpha_n}f^*{}_{\alpha_1\ldots\alpha_n}(0)| \leqq \{\sum |f^*{}_{\alpha_1\ldots\alpha_n}(0)|^2\}^{1/2}$$

Now we return to the proof of Theorem 4. We shall prove the validity of (2.9). To do this we fix the geodesic and the point x on it. If now we consider a neighborhood of x, and normal coördinates with this point as origin, we may apply the lemma to the scalar f in this neighborhood. In particular, from (2.14) we infer that

$$|F^{(n)}(s)| \leqq \{\sum |f^*{}_{\alpha_1\ldots\alpha_n}(0)|^2\}^{1/2} \qquad (n = 1, 2, \cdots).$$

But

$$D_n(f,x) = \psi^{\alpha_1\beta_1}(0)\cdots\psi^{\alpha_n\beta_n}(0)f^*{}_{\alpha_1\ldots\alpha_n}(0)f^*{}_{\beta_1\ldots\beta_n}(0)$$
$$= \{\sum |f^*{}_{\alpha_1\ldots\alpha_n}(0)|^2\}^{1/2}$$

because of (2.3). Hence $|F^{(n)}(s)| \leqq D_n(f,x)$. Thus the proof is complete.

Remark. In Theorems 3 and 4 we may replace the assumption that x_0 is an interior point of R by the requirement that it be a boundary point of R which is the vertex of a conical sub-region of R filled out entirely by geodesic arcs issuing from x_0, the condition b) being in this case replaced by the condition

b′) $\lim D_n(f,x) = 0$ as $x \to x_0$ along a path in R.

3. In this section we shall obtain three theorems in which restrictions

are placed, not on the quantities $D_n(f,x)$ themselves, but on the integrals of these expressions.

THEOREM 5. *If f is defined in a connected domain R in V_k, and $D_n(f,x)$ has the meaning ascribed to it in either of the two preceding theorems, then f vanishes identically provided that conditions* a)–c) *below are fulfilled.*

a) $\int_R D_{n+1}(f,x)\,dv \leqq m_n \qquad (n = -1, 0, 1, \cdots)$

b) $f = 0$ in a neighborhood of a point x_0 in R

c) $\sum m_n^{-1/n} = \infty$,

(m_{-1} is a positive constant, and $dv = \sqrt{g}\,dx^1 \cdots dx^k$ is the volume element in V_k).

Proof. Consider the non-negative function

$$\Phi(x) = \sum_0^\infty \frac{1}{2^n m_n} D_{n+1}(f,x).$$

From a) we see that

$$\int_R \Phi\,dv \leqq 2.$$

Let S_0 be a geodesic sphere with center at x_0, lying in R, and entirely covered by a system of normal coördinates with origin at x_0. Since in such a sphere $\sqrt{g} \geqq c > 0$ we conclude that the integral

$$\int_{S_0} \Phi\,dy^1 \cdots dy^k$$

is finite, y^i being normal coördinates. Now we may evaluate this integral by polar coördinates,

$$\int_{S_0} \Phi\,dy^1 \cdots dy^k = \int_{W_k} \left(\int_0^{r_0} s^{k-1}\Phi\,ds \right) d\omega$$

where r_0 is the radius of S_0, and W_k, $d\omega$ denote the surface of the k-dimensional unit sphere and its surface element, respectively. By Fubini's theorem we conclude that

$$\int_0^{r_0} s^{k-1}\Phi\,ds$$

is finite for almost all geodesics issuing from x_0. But f, and therefore Φ, vanishes in the neighborhood of x_0. Consequently the integral remains finite with the factor s^{k-1} of the integrand suppressed.

It results from these considerations that if we define $F(s)$ as in the preceding work then since $|F^{(n)}(s)| \leqq D_n(f,x)$ and $F^{(n)}(0) = 0$ we have

$$|F^{(n)}(s)| = |\int_0^s F^{(n+1)}(s)\,ds| \leqq \int_0^{r_0} D_{n+1}(f,x)\,ds$$
$$|F^{(n)}(s)| \leqq C^n m_n \qquad 0 \leqq s \leqq r_0$$

where C is a constant depending on f and the particular geodesic. By Theorem D, $F(s)$ must be zero, so that f vanishes along almost all geodesics through x_0 in S_0. It then vanishes identically, for it is continuous. The conclusion of the theorem is then reached as before.

The stringency of condition b) can be removed if we alter c) somewhat. We shall do this first for Euclidean space.

THEOREM 6. *Let f be defined throughout E_k, and let x_0 be an arbitrary point. Then the conditions*

a) $\int_{E_k} D_n(f,x)\,dv \leqq m_n$

b) $D_n(f,x_0) = 0$

c) $\sum (m_n + \cdots + m_{n+k-1})^{-1/n} = \infty$

imply that $f = 0$.

Proof. Let

$$\Phi(x) = \sum_0^\infty \frac{1}{2^n m_n} D_n(f,x).$$

Then by a)

$$\int_{E_k} \Phi\,dv \leqq 2.$$

If we adopt spherical coördinates about the point x_0 and denote by s the distance measured along a half-line issuing from x_0, then we conclude by Fubini's theorem, in a manner similar to that of the previous proof, that for almost all half-lines emanating from x_0 the integral

$$\int_0^\infty s^{k-1}\Phi\,ds$$

is finite, and therefore, for some constant A depending on f and the particular line

$$\int_0^\infty s^{k-1} D_n(f,x(s))\,ds \leqq A^n m_n.$$

Now, in the notation used in proving Theorem 1, $F(s) = f(x(s))$ and $|F^{(n)}(s)| \leqq D_n(f,x(s))$. Thus

$$(3.1) \qquad \int_0^\infty |F^{(n)}(s)|\,s^{k-1}ds \leqq A^n m_n$$

$$F^{(n)}(0) = 0. \tag{3.2}$$

It will be proved in the lemma which follows that (3.1) and (3.2), together with condition c) imply that $F(s) \equiv 0$ $(0 \leqq s < \infty)$. Since f is continuous in E_k, this implies that $f = 0$, thus completing the proof. We turn now to the lemma.

LEMMA 2. *Let F be a real function of class C^∞ on the infinite interval $(0, \infty)$, and let the relations* (3.1), (3.2) *above, and condition* c) *of Theorem 6 be fulfilled for an integer $k \geqq 1$. Then $F \equiv 0$.*

Proof. $F^{(n)}$ is absolutely integrable, by (3.1). Therefore, for complex z with positive real part $(R(z) > 0)$ the integral

$$H(z) = \int_0^\infty e^{-zs} F(s)\,ds \tag{3.3}$$

exists and defines an analytic function. Integrating by parts repeatedly and using (3.2) we obtain

$$z^n H(z) = \int_0^\infty e^{-zs} F^{(n)}(s)\,ds.$$

Now we write $H_n(z) = z^n H(z)$. Then

$$\frac{d^{k-1} H_n(z)}{dz^{k-1}} = (-1)^{k-1} \int_0^\infty s^{k-1} e^{-zs} F^{(n)}(s)\,ds$$

and if $R(z) > 0$ we infer from (3.1) that

$$\left| \frac{d^{k-1} H_n(z)}{dz^{k-1}} \right| \leqq A^n m_n. \tag{3.4}$$

For a fixed z, $R(z) > 0$, we define

$$\phi(t) = H_n((1-t)z + t) \qquad 0 \leqq t \leqq 1. \tag{3.5}$$

Integration by parts, applied $k-2$ times to the equation

$$\phi(1) - \phi(0) = \int_0^1 \phi'(t)\,dt$$

yields

$$\phi(0) = \sum_{\nu=0}^{k-2} (-1)^\nu \frac{1}{\nu!} \phi^{(\nu)}(1) + \frac{(-1)^{k-1}}{(k-2)!} \int_0^1 t^{k-2} \phi^{(k-1)}(t)\,dt.$$

Hence

$$H_n(z) = P_n(z) + \frac{(-1)^{k-1}}{(k-2)!} \int_0^1 t^{k-2} \phi^{(k-1)}(t)\,dt \tag{3.6}$$

where

$$P_n(z) = \sum_{\nu=0}^{k-2} \frac{(z-1)^\nu}{\nu!} H_n^{(\nu)}(1) \tag{3.7}$$

and

$$\phi^{k-1}(t) = (-1)^{k-1}(z-1)^{k-1} H_n^{(k-1)}((1-t)z+t).$$

Combining (3.4) and (3.6) we find

$$| H_n(z) - P_n(z) | \leqq \frac{|z-1|^{k-1}}{(k-2)!} A^n m_n. \tag{3.8}$$

$P_n(z)$ as a function of n is a polynomial of degree $\leqq k-2$. Hence, differencing with respect to n,

$$\sum_{\nu=0}^{k-1} (-1)^\nu \binom{k-1}{\nu} P_{n-\nu}(z) = 0 \qquad (n = k-1, k, \cdots). \tag{3.9}$$

Now

$$\sum_{\nu=0}^{k-1} (-1)^\nu \binom{k-1}{\nu} H_{n-\nu}(z) = H(z) \sum_{\nu=0}^{k-1} (-1)^\nu \binom{k-1}{\nu} z^{n-\nu} \tag{3.10}$$
$$= H(z) z^{n-k+1} (z-1)^{k-1}.$$

Thus, combining (3.8), (3.9), (3.10)

$$| H(z) z^{n-k+1} (z-1)^{k-1} | \leqq \sum_{\nu=0}^{k-1} \binom{k-1}{\nu} \frac{|z-1|^{k-1}}{(k-2)!} A^{n-\nu} m_{n-\nu}.$$

From this inequality we obtain, making the unimportant assumption that $A \geqq 1$, the inequalities

$$| z^n H(z) | \leqq A^{n+k-1}(m_n + \cdots + m_{n+k-1}). \tag{3.11}$$

It then follows by condition c) and a lemma of Carleman [6] (which he uses for the proof of Theorem D) that the analytic function $H(z)$ is identically zero. Then $F \equiv 0$ also, as was to be proved.

Before going on to the next theorem we shall prove another lemma.

LEMMA 3. *Let $f(x)$ be of class C^∞ on $(0,1)$, and suppose that*

a) $\int_0^1 | f^{(n)}(x) | x^{k-1} dx \leqq A^n m_n$

b) $f^{(n)}(0) = 0$

c) $m_1^{-1} \geqq m_2^{-1/2} \geqq \cdots \geqq m_n^{-1/n} \geqq \cdots \to 0$

d) $\sum (m_{n+k-1})^{-1/n} = \infty$

---

[6] T. Carleman, *op. cit.*, p. 20, Corollary. The statement is that if $\Phi(z)$ is analytic in the half plane $R(z) \geqq a \geqq 0$, and if $| \Phi(z) z^{\lambda_n} | \leqq \beta_n^{\lambda_n}$, where $0 < \lambda_0 < \lambda_1 < \cdots < \lambda_n < \cdots \to \infty$ and $\beta_n > 0$, then the divergence of the series

$$\Sigma \frac{\lambda_n - \lambda_{n-1}}{\beta_n}$$

implies $\Phi \equiv 0$.

where k is an integer $\geqq 1$ and A is a constant (which we may assume to be $\geqq 1$). *Then $f \equiv 0$.*

Proof. We shall reduce the problem to that of Lemma 2. Putting

$$x = \frac{t}{1+t} \qquad g(t) = f(x) \qquad 0 \leqq t < \infty$$

we calculate the derivatives of $g(t)$:

$$(3.12) \quad g^{(n)}(t) = \frac{a_{n,n}}{(1+t)^{2n}} f^{(n)}(x) + \frac{a_{n,n-1}}{(1+t)^{2n-1}} f^{(n-1)}(x) + \cdots + \frac{a_{n,1}}{(1+t)^{n+1}} f^{(1)}(x)$$

where the $a_{n,\nu}$ are constants, with $a_{n,n} = 1$. Thus, using a)

$$(3.13) \qquad \int_0^\infty \frac{|g^{(n)}(t)|}{(1+t)^{k+1}} t^{k-1} dt \leqq A^n \sum_0^n |a_{n,\nu}| m_\nu.$$

If we differentiate (3.12) we obtain the recursion formula

$$(3.14) \qquad a_{n+1,\nu+1} = a_{n,\nu} - (n+\nu+1) a_{n,\nu+1}, \qquad (\nu = 0, \cdots, n)$$

which is correct in all cases if we define $a_{n,0} = 0$, $a_{n,\nu} = 0$ if $\nu > n$. Therefore

$$(3.15) \qquad |a_{n+1,\nu+1}| \leqq |a_{n,\nu}| + 2n |a_{n,\nu+1}|.$$

Consider now the expression

$$P_n(\xi) = \xi(\xi+2)\cdots(\xi+2n-2) = \sum_0^n b_{n,\nu}\xi^\nu.$$

$P_n(\xi)$ is a polynomial of degree n in ξ. Clearly $b_{n,\nu} \geqq 0$ and

$$(3.16) \qquad P_n(\xi) \leqq (\xi + 2n)^n \qquad \xi \geqq 0.$$

It is easily verified that $b_{n,n} = 1$ and that

$$(3.17) \qquad b_{n+1,\nu+1} = b_{n,\nu} + 2n b_{n,\nu+1}$$

where we again adopt the conventions $b_{n,0} = 0$, $b_{n,\nu} = 0$ if $\nu > n$. Since $a_{11} = b_{11} = 1$ a comparison of (3.15) and (3.17) shows that

$$(3.18) \qquad |a_{n,\nu}| \leqq b_{n,\nu}.$$

Now let us write $\alpha_n = m_n^{1/n}$. By c) the α's have the property

$$(3.19) \qquad \alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n \leqq \cdots \to \infty.$$

From (3.13), (3.18), (3.19) and (3.16) we conclude

$$\int_0^\infty \frac{|g^{(n)}(t)|}{(1+t)^{k+1}} t^{k-1} dt \leqq A^n(\alpha_n + 2n)^n. \tag{3.20}$$

Finally, we set

$$h(t) = \frac{g(t)}{(1+t)^{k+1}}.$$

Then

$$h^{(n)}(t) = \sum_0^n (-1)^\nu \binom{n}{\nu} (k+1) \cdots (k+\nu) \frac{g^{(n-\nu)}(t)}{(1+t)^{k+1+\nu}}$$

and so

$$|h^{(n)}(t)| \leqq \sum_{\nu=0}^n \binom{n}{\nu} (k+n)^\nu \frac{|g^{(n-\nu)}(t)|}{(1+t)^{k+1}}.$$

From this result and (3.19), (3.20) we get

$$\int_0^\infty |h^{(n)}(t)| t^{k-1} dt \leqq A^n(\alpha_n + 3n + k)^n$$

or, for a suitable constant B (depending only on A and k),

$$\int_0^\infty |h^{(n)}(t)| t^{k-1} dt \leqq B^n(\alpha_n + 4n)^n. \tag{3.21}$$

Since $h^{(n)}(0) = 0$ the conclusion $h(t) \equiv 0$ will follow from Lemma 2 if we can prove that, writing $\beta_n = \alpha_n + 4n$, we have

$$\sum (\beta_n{}^n + \cdots + \beta_{n+k-1}^{n+k-1})^{-1/n} = \infty. \tag{3.22}$$

Now $\beta_1 \leqq \beta_2 \leqq \cdots \leqq \beta_n \leqq \cdots \to \infty$. Therefore the series in (3.22) will diverge if

$$\sum (\beta_{n+k-1})^{-1-(k-1)/n} = \infty. \tag{3.23}$$

In proving this we shall distinguish two cases. Suppose first that for all sufficiently large values of n, $\alpha_n \geqq n$, $\beta_n \leqq 5\alpha_n$. In this case the series (3.23) will diverge if

$$\sum (\alpha_{n+k-1})^{-1-(k-1)/n} = \infty.$$

But this is true because of condition d) of the lemma and the definition of α_n. The alternative situation is that in which there exists a strictly increasing sequence $\{n_\nu\}$ of integers such that $\alpha_{n_\nu} < n_\nu$ ($\nu = 1, 2, \cdots$). Thus we have

$$\beta_n \leqq 5n_\nu \quad \text{if } n_{\nu-1} < n \leqq n_\nu. \tag{3.24}$$

Now with the omission of at most some early terms the series (3.23) may be written

$$\sum (\beta_n)^{-n/(n-k+1)} = \sum_{\nu} \sum_{n=n_{\nu-1}+1} (\beta_n)^{-n/(n-k+1)}$$

and so, by (3. 24), it diverges if the series

$$\sum_{\nu} \sum_{n=n_{\nu-1}+1} (n_\nu)^{-n/(n-k+1)} \tag{3.25}$$

is divergent.

In proving the divergence of (3. 25) we may obviously assume that n_ν tends to infinity as rapidly as we please, for this reduces the size of the partial sums. In particular if we can show that for some constant $\mathcal{E}_0 > 0$ and each integer $m \geqq k$ it is possible to determine an integer $N = N(m)$ so large that

$$\sum_{n=m}^{N} N^{-n/n-k+1} \geqq \mathcal{E}_0$$

then we see that with $n_{\nu-1}$ chosen, n_ν may be selected so as to contribute at least $\mathcal{E}_0$ to the sum of the series, which must then diverge. Now for fixed $m > k$

$$\lim \sum_{n=k}^{m} N^{-n/n-k+1} = 0 \text{ as } N \to \infty.$$

Hence we need only prove that

$$\liminf \sum_{n=k}^{N} N^{-n/n-k+1} > 0 \text{ as } N \to \infty. \tag{3.26}$$

Now

$$\sum_{n=k}^{N} N^{-n/n-k+1} \geqq \int_{k}^{N+1} N^{-t/t-k+1} dt.$$

Making the change of variable $u = \log N/t - k + 1$ we obtain, as soon as $\log N > 1$,

$$\sum_{n=k}^{N} N^{-n/n-k+1} \geqq \frac{\log N}{N} \int_{a/N-k+2}^{1} e^{-(k-1)u} \frac{du}{u^2} \quad (a = \log N)$$
$$\geqq \frac{\log N}{N} e^{-(k-1)} [(N - k + 2)/\log N - 1]$$

and so

$$\liminf \sum_{n=k}^{N} N^{-n/n-k+1} \geqq e^{-(k-1)} > 0 \text{ as } N \to \infty.$$

This completes the proof of the lemma. It is to be used in the proof of the following theorem.

THEOREM 7. *Let f be defined in a connected domain R in V_k. Then if x_0 is a point in R and $D_n(f, x)$ has the significance ascribed to it in either Theorem 3 or Theorem 4, the conditions*

a) $\int_R D_n(f, x)\, dv \leqq m_n$

b) $D_n(f, x_0) = 0$

together with conditions c), d) *of Lemma* 3, *imply that* $f \equiv 0$.

Proof. We may show, by considerations similar to those involved in the proof of Theorem 5, that for a geodesic sphere with center at x_0 and radius r_0 the inequalities

$$\int_0^{r_0} | F^{(n)}(s) |\, s^{k-1} ds \leqq C^n m_n$$

are satisfied for almost all geodesics through x_0. By a trivial transformation we may assume that $r_0 = 1$. Then since $F^{(n)}(0) = 0$ the previous lemma assures us that $F(s) \equiv 0$ for almost all geodesics. The remainder of the proof then goes as before.

Part II.

4. In this section we shall present two theorems about functions defined in the whole Euclidean space E_k. They differ from the theorems of Part I in two respects. Instead of the quantities $D_n(f, x)$ defined by (1. 1) we consider differential operators applied to f. These need not involve all the partial derivatives of f. The second difference lies in the fact that we replace the single point x_0 by a certain type of point set. For this we lay down the following definition:

Definition. *A point set U contained in a region R of E_k will be called a set of analytic determination, or simply a set of determination, if a function ϕ which is analytic in R and which has the value zero at all points of U must vanish identically in R.*

For the first theorem the differential operators are defined in the following manner. Let p be a fixed positive integer. For each positive integer n let $P_n(\xi_1, \cdots, \xi_k)$ be a homogeneous polynomial in $\xi_1, \cdots, \xi_k$ of degree np, with real coefficients. We define $P_0(\xi_1, \cdots, \xi_k) = 1$. Then if f and all its derivatives are defined in E_k the differential operators

$$P_n\left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_k}\right)$$

have an obvious meaning when applied to f, the resulting being a function defined in E_k. We use the notation

$$P_n\left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_k}\right) f = L_n(f, x). \tag{4. 1}$$

THEOREM 8. *Let f be defined in E_k. With a certain fixed Cartesian coördinate system $(x_1, \cdots, x_k)$ let $f(x) = 0$ except when $x_i > 0$ $(i = 1, \cdots, k)$, and let U be a set of determination contained in the region $x_i > 0$ $(i = 1, \cdots, k)$. Then if*

$$\text{a)} \quad | L_n(f, x) | \leqq m_n, \ x \text{ in } E_k$$

$$\text{b)} \quad \sum | P_n(\xi) |^{1/np} (m_n)^{-1/np} = \infty, \ \xi \text{ in } U$$

the function f vanishes identically.

Proof. We shall first carry out the proof under the additional assumption that all the partial derivatives of f are bounded in E_k. (f itself is bounded, because of a).

We form the function

$$F(z) \equiv F(z_1, \cdots, z_k) = \int_0^\infty \cdots \int_0^\infty e^{-(z_1 t_1 + \cdots + z_k t_k)} f(t) \, dt_1 \cdots dt_k.$$

Since f is bounded F is analytic if $R(z_i) > 0$, $i = 1, \cdots, k$. Under our temporary assumption we may integrate by parts; since $f(x) = 0$ if any $x_i = 0$ (and similarly for the derivatives) we obtain, for any non-negative integers $\rho_1, \cdots, \rho_k$

$$z_1^{\rho_1} \cdots z_k^{\rho_k} F(z) = \int_0^\infty \cdots \int_0^\infty e^{-(z_1 t_1 + \cdots + z_k t_k)} \frac{\partial^{\rho_1 + \cdots + \rho_k}}{\partial t_1^{\rho_1} \cdots \partial t_k^{\rho_k}} f(t) \, dt_1 \cdots dt_k$$

$$P_n(z) F(z) = \int_0^\infty \cdots \int_0^\infty e^{-(z_1 t_1 + \cdots + z_k t_k)} L_n(f, t) \, dt_1 \cdots dt_k$$

and if $R(z_i) = x_i > 0$ it follows that

$$| P_n(z) F(z) | \leqq \frac{m_n}{x_1 \cdots x_k}.$$

Now let $(\xi_1, \cdots, \xi_k)$ be any point in U, and place $z_i = \xi_i w$, where $R(w) > 0$. Then by the homogeneity of P_n we obtain

$$| w^{np} P_n(\xi) F(\xi_1 w, \cdots, \xi_k w) | \leqq \frac{m_n}{[R(w)]^k \xi_1 \cdots \xi_k}.$$

For fixed ξ define $\Phi(w) = F(\xi_1 w, \cdots, \xi_k w) \cdot \Phi(w)$ is analytic if $R(w) > 0$. If

$$R(w) \{\xi_1 \cdots \xi_k\}^{1/k} \geqq 1$$

we have

$$| w^{np} \Phi(w) | \leqq \frac{m_n}{P_n(\xi)}. \tag{4.2}$$

But by condition b) it follows from a result of Carleman already referred to,[7]

[7] Cf. footnote 6.

that $\Phi(w) \equiv 0$. In this way we see that the analytic function F vanishes at all points of U, and therefore $F \equiv 0$. This implies that $f \equiv 0$.

We have now to free ourselves of the assumption of boundedness of the partial derivatives of f. We do this by approximating the function f by a function whose derivatives are bounded.

By a general theorem [8] it is known that if K is an absolutely integrable function defined over E_k, with

$$\int_{E_k} K(x_1, \cdots, x_k)\,dx_1 \cdots dx_k = 1 \tag{4.3}$$

and if f is bounded and continuous, then for the function $(\lambda > 0)$

$$f_\lambda(x) = \int_{E_k} f\left(x_1 + \frac{t_1}{\lambda}, \cdots, x_k + \frac{t_k}{\lambda}\right) K(t_1, \cdots, t_k)\,dt_1 \cdots dt_k \tag{4.4a}$$

$$f_\lambda(x) = \lambda^k \int_{E_k} f(t_1, \cdots, t_k) K(\lambda(t_1 - x_1), \cdots, \lambda(t_k - x_k))\,dt_1 \cdots dt_k \tag{4.4b}$$

we have

$$\lim f_\lambda(x) = f(x) \quad \text{as } \lambda \to \infty. \tag{4.5}$$

We shall require of K that it be non-negative, possess derivatives of all orders, and that it vanish outside a fixed, finite hypercube in E_k. Then it is permissible to differentiate under the integral in (4.4a), and we obtain

$$L_n(f_\lambda, x) = \int_{E_k} L_n\left(f, x + \frac{t}{\lambda}\right) K(t)\,dt \tag{4.6}$$

where we are using an obvious abbreviation in notation. Consequently, by condition a), together with (4.3) and the fact that K is non-negative, we conclude

$$|\, L_n(f_\lambda, x) \,| \leqq m_n.$$

Since K vanishes outside a finite region in E_k, and f vanishes except when $x_i > 0$, we easily infer from (4.4) that for any fixed $\lambda > 0$, f_λ vanishes except when $x_i > b_i$, where $b_1, \cdots, b_k$ are constants depending on λ. Finally, the partial derivatives of f_λ are bounded in E_k. To see this we make use of (4.4b):

$$\begin{aligned} \frac{\partial^{\rho_1 + \cdots + \rho_k}}{\partial x_1^{\rho_1} \cdots \partial x_k^{\rho_k}} f_\lambda(x) &= (-\lambda)^{\rho_1 + \cdots + \rho_k} \lambda^k \int_{E_k} f(t) \left[\frac{\partial^{\rho_1 + \cdots + \rho_k}}{\partial u_1^{\rho_1} \cdots \partial u_k^{\rho_k}} K(u) \right]_{u = \lambda(t-x)} dt \\ &= (-\lambda)^{\rho_1 + \cdots + \rho_k} \int_{E_k} f\left(x + \frac{t}{\lambda}\right) \frac{\partial^{\rho_1 + \cdots + \rho_k}}{\partial t_1^{\rho_1} \cdots \partial t_k^{\rho_k}} K(t)\,dt. \end{aligned}$$

Since the derivatives of K vanish outside a finite region, and f is bounded, the result follows.

---

[8] S. Bochner, *Vorlesungen über Fouriersche Integrale*, Leipzig (1932), p. 191.

But now we can conclude from the first part of the proof, after a trivial change of coördinates, that $f_\lambda \equiv 0$. By (4.5) it follows that $f \equiv 0$. The proof is thus complete.

For the next theorem we introduce the Laplace operator and its iterates.

$$\Delta f(x) = \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_k^2}$$
$$\Delta^n f(x) = \Delta(\Delta^{n-1} f(x)).$$

As usual we adopt the convention $\Delta^0 f = f$.

THEOREM 9. *Let f be defined in E_k, and let U be a set of determination in E_k. Then the conditions*

a) $|\Delta^n f(x)| \leqq m_n$, x *in* E_k

b) $\Delta^n f(\xi) = 0$, ξ *in* U

c) $\sum m_n^{-1/n} = \infty$

imply that $f \equiv 0$.

Proof. We again make the temporary assumption that the partial derivatives of f are bounded in E_k. For $\alpha \geqq 0$ we define

(4.7a) $$F(x_1, \cdots, x_k, \alpha) = c_k \int_{E_k} f(x_1 + \sqrt{\alpha}\, t_1, \cdots, x_k + \sqrt{\alpha}\, t_k)\, e^{-(t_1^2 + \cdots + t_k^2)/4} dt_1 \cdots dt_k$$

(4.7b) $$F(x_1, \cdots, x_k, \alpha) = c_k \alpha^{-k/2} \int_{E_k} f(t_1, \cdots, t_k)\, e^{-[(t_1 - x_1)^2 + \cdots + (t_k - x_k)^2]/\alpha} dt_1 \cdots dt_k$$

where $c_k = 2^{-k}\pi^{-k/2}$ is a normalizing factor. $F(x_1, \cdots, x_k, \alpha) \equiv F(x, \alpha)$ may be regarded as a transform of $f(x)$. If we write

$$F(x, \alpha) = T_\alpha(f, x)$$

then because of the formula

$$c_k \int_{E_k} e^{-(t_1^2 + \cdots + t_k^2)/4} dt_1 \cdots dt_k = 1$$

it is clear that $T_0(f, x) = f(x)$ and

(4.8) $$|T_\alpha(f, x)| \leqq \max |f|.$$

It is well known that [9]

$$(4.9) \qquad \lim F(x, \alpha) = f(x) \quad \text{as} \quad \alpha \to \infty$$

the convergence being uniform in any finite portion of E_k.

Under our temporary assumption we may calculate the derivative of $F(x, \alpha)$ with respect to α by differentiating under the integral in (4.7a). If we do this, and then integrate the resulting equation by parts we obtain

$$\frac{\partial F(x, \alpha)}{\partial \alpha} = T_\alpha(\Delta f, x).$$

Consequently

$$(4.10) \qquad \frac{\partial^n F(x, \alpha)}{\partial \alpha^n} = T_\alpha(\Delta^n f, x).$$

Therefore, by (4.8) and condition a)

$$\left| \frac{\partial^n F(x, \alpha)}{\partial \alpha^n} \right| \leqq |\Delta^n f| \leqq m_n, \quad x \text{ in } E_k, \quad \alpha \geqq 0$$

while by condition b)

$$\left(\frac{\partial^n F(\xi, \alpha)}{\partial \alpha^n} \right)_{\alpha=0} = \Delta^n f(\xi) = 0, \quad \xi \text{ in } U.$$

By Theorem D it follows that $F(\xi, \alpha) = 0$ if $\alpha \geqq 0$ and ξ is in U. But for fixed $\alpha > 0$, $F(x, \alpha)$ is analytic in the entire space E_k. This is proved [10] by making use of (4.7b). Consequently, because of the nature of the set U, $F(x, \alpha) = 0$ for each x in E_k, if $\alpha > 0$. By (4.9) we conclude $f \equiv 0$.

Without the assumption of boundedness of the partial derivatives of f the whole difficulty lies in the establishment of (4.10). In order to meet this difficulty we define $f_\lambda(x)$ by (4.4), where K has the same properties as were required before. Then we know that the partial derivatives of f_λ $(\lambda > 0)$ with respect to the x's are bounded in E_k, and that

$$(4.11) \qquad |\Delta^n f_\lambda(x)| \leqq m_n, \quad x \text{ in } E_k, \quad \lambda > 0.$$

From (4.6) it follows that

$$(4.12) \qquad \lim \Delta^n f_\lambda(x) = \Delta^n f(x) \quad \text{as} \quad \lambda \to \infty$$

---

[9] This is in fact a special case of the theorem referred to in footnote 8. For the case $k = 1$ see E. Borel, *Leçons sur les Fonctions de Variables Reélles*, Paris (1928), pp. 52-53.

[10] The method used by Borel, *op. cit.*, pp. 53-54 for the case $k = 1$ works equally well in general. It consists in defining the integral for complex values of the variables and showing that the resulting function is analytic.

and it is easily verified that the convergence is uniform over each finite position of E_k. Now by the previous work, if

$$F_\lambda(x, \alpha) = T_\alpha(f_\lambda, x)$$

then

$$\frac{\partial^n F_\lambda(x, \alpha)}{\partial \alpha^n} = T_\alpha(\Delta^n f_\lambda, x). \tag{4.13}$$

It is readily proved, using condition a) and (4.11), (4.12), that

$$\lim T_\alpha(\Delta^n f_\lambda, x) = T_\alpha(\Delta^n f, x) \quad \text{as } \lambda \to \infty.$$

This follows from (4.7b), together with the fact that f, f_λ are bounded, and that $f_\lambda \to f$ uniformly over each finite portion of E_k.

The passage to the limit $\lambda \to \infty$ in (4.13) then establishes (4.10) and completes the proof.

5. In this concluding section we are concerned with the sphere

$$V_k: \quad x_1^2 + \cdots + x^2_{k+1} = 1 \qquad (k \geqq 2)$$

in E_{k+1}. It is an analytic Riemannian space of constant positive curvature, and as such the Laplace operator Δ is defined for it.[11]

We may define a set of determination U on V_k in the same way as it was done in § 4. A function ϕ is analytic over a portion R of V_k if it is analytic in each of the allowable coördinate systems covering R (any two coördinate systems being connected by analytic transformations).

Points of V_k will be denoted by $P, Q, \cdots$.

THEOREM 10. *Let f be defined in V_k $(k \geqq 2)$, and let U be a set of determination in V_k. Then the conditions*

a) $|\Delta^n f(P)| \leqq m_n$, P *in* V_k

b) $\Delta^n f(P_0) = 0$, P_0 *in* U

c) $\sum m_n^{-1/n} = \infty$

imply that $f \equiv 0$.

The proof depends upon the properties of certain expansions in series of ultra-spherical functions. These expansions reduce to the ordinary Laplace series of spherical harmonics when $k = 2$.

[11] The invariant form of the operator is

$$\Delta f = g^{ij} f_{,i,j}.$$

See O. Veblen, *Invariants of Quadratic Differential Forms*, Cambridge University Press (1927), p. 63.

We shall introduce coördinates in V_k by using spherical coördinates in E_{k+1}. We put

$$
(5.1)\qquad
\begin{aligned}
x_1 &= r\cos\theta_1\\
x_2 &= r\sin\theta_1\cos\theta_2\\
&\cdots\cdots\\
x_k &= r\sin\theta_1\cdots\sin\theta_{k-1}\cos\theta_k\\
x_{k+1} &= r\sin\theta_1\cdots\sin\theta_k.
\end{aligned}
$$

Then it may be found that if [12]

[12] See E. Heine, *Handbuch der Kugelfunktionen*, Band I, 2nd ed., Berlin (1878), pp. 460-461.

$$
\begin{aligned}
v &= r^k\sin^{k-1}\theta_1\sin^{k-2}\theta_2\cdots\sin\theta_{k-1}\\
u_1 &= r^2\\
u_2 &= r^2\sin^2\theta_1\\
&\cdots\cdots\\
u_k &= r^2\sin^2\theta_1\cdots\sin^2\theta_{k-1}
\end{aligned}
$$

and F is defined in E_{k+1}, then

$$\sum_{i=1}^{k+1}\frac{\partial^2 F}{\partial x_i^2}=\frac{1}{v}\left[\frac{\partial}{\partial r}\left(\frac{v}{1}\frac{\partial F}{\partial r}\right)+\frac{\partial}{\partial\theta_1}\left(\frac{v}{u_1}\frac{\partial F}{\partial\theta_1}\right)+\cdots+\frac{\partial}{\partial\theta_k}\left(\frac{v}{u_k}\frac{\partial F}{\partial\theta_k}\right)\right]$$

Hence we have, for a function f defined in V_k

$$(5.2)\qquad \Delta f=\sum_{i=1}^{k}\frac{w_i}{\sin^{k-1}\theta_i}\frac{\partial}{\partial\theta_i}\left(\sin^{k-1}\theta_i\frac{\partial f}{\partial\theta_i}\right)$$

where $w_1=1$, $w_i=(\sin\theta_1\cdots\sin\theta_{i-1})^{-2}$, $i=2,\cdots,k$.

Now we put

$$(5.3)\qquad \nu=(k-1)/2$$

and consider the polynomials $P_n^{(\nu)}(x)$ defined by the expansion [13]

$$(5.4)\qquad (1-2rx+r^2)^{-\nu}=\sum_0^\infty r^nP_n^{(\nu)}(x).$$

For $k=2$, $\nu=1/2$ and $P_n^{(\nu)}(x)$ is the Legendre polynomial of order n. It is known that if P, Q are two points on the sphere V_k and γ is the angle $(0\leqq\gamma\leqq\pi)$ between the lines joining these points to the center of the sphere (so that γ is the geodesic distance from P to Q), then for the function

$$(5.5)\qquad F_n(P,Q)=P_n^{(\nu)}(\cos\gamma)$$

[13] A systematic study of the functions $P_n^{(\nu)}(x)$ may be found in N. Nielsen, *Théorie des Fonctions Metasphériques*, Paris (1911), especially Chapter VII. The expansion (5.4) is given on p. 98. The notation of Heine, *op. cit.*, pp. 451-460, is slightly different from that of Nielsen.

we have [14]

$$\Delta F_n = -n(n+2\nu)F_n. \tag{5.6}$$

Since F_n is symmetric in P and Q it does not matter which point is regarded as variable in evaluating the left member of (5.6).

As a consequence of (5.6) and Green's theorem we also have the important relation

$$\int_{V_k} F_n(P,Q)\Delta f(Q)dQ = -n(n+2\nu)\int_{V_k} F_n(P,Q)f(Q)dQ \tag{5.7}$$

where dQ denotes the volume element in the space V_k.

With a function f defined in V_k we associate the series [15]

$$f(P) \sim \sum_0^\infty a_n(P,f) \tag{5.8}$$

where

$$a_n(P,f) = \frac{(n+\nu)\Gamma(\nu)}{2\pi^{\nu+1}}\int_{V_k} F_n(P,Q)f(Q)dQ. \tag{5.9}$$

Concerning this series we assert:

LEMMA 4. *If f is continuous on V_k, and $H_n^{(p)}(P)$ denotes the Cesàro means of order p of the series (5.8), then*

$$\lim_{n\to\infty} H_n^{(p)}(P) = f(P) \quad \text{if } p > \nu$$

$$|H_n^{(p)}(P)| \leqq \max|f| \quad \text{if } p \geqq 2\nu+1 = k.$$

Proof. In proving the first assertion we may assume that the coördinates (5.1) are chosen so that P is the point $x_1 = 1$, $x_i = 0$, $i > 1$. Then $\gamma = \theta_1$ and

$$\int_{V_k} F_n(P,Q)f(Q)dQ = \int_0^\pi d\theta_1 \cdots \int_0^\pi d\theta_{k-1}\int_\pi^\pi P_n^{(\nu)}(\cos\theta_1)\sin^{k-1}\theta_1 f(Q)\omega d\theta_k$$

where $\omega = \sin^{k-2}\theta_2 \cdots \sin\theta_{k-1}$ and Q has the coördinates $(\theta_1,\cdots,\theta_k)$. From this result and the formula [16]

$$P_n^{(\nu)}(1) = \frac{\Gamma(n+2\nu)}{\Gamma(n+1)\Gamma(2\nu)}$$

[14] Heine, *op. cit.*, p. 461.

[15] This series arises naturally through the expansion of the integrand in Poisson's integral when the latter is used to solve the Dirichlet problem for the sphere V_k.

[16] Nielsen, *op. cit.*, p. 95.

it is seen that if we put

$$(5.10) \quad \Phi(\cos\theta_1) = \frac{\Gamma(\nu+1/2)\Gamma(1/2)}{2\pi^{\nu+1}} \int_0^\pi d\theta_2 \cdots \int_0^\pi d\theta_{k-1} \int_{-\pi}^{\pi} f(Q)\,\omega d\theta_k$$

then the series (5.8) becomes the formal expansion of $\Phi(\cos\theta)$ in a series of the form

$$(5.11) \qquad \Phi(\cos\theta) \sim \sum_0^\infty c_n P_n^{(\nu)}(\cos\theta)$$

for the point $\theta = 0$, the coefficients c_n being given by the formula [17]

$$(5.12) \quad c_n = \frac{(n+\nu)\Gamma(\nu)}{\Gamma(\nu+1/2)\Gamma(1/2)} \frac{\Gamma(n+1)\Gamma(2\nu)}{\Gamma(n+2\nu)} \int_0^\pi \Phi(\cos\theta) P_n^{(\nu)}(\cos\theta)\sin^{2\nu}\theta d\theta.$$

$\Phi(\cos\theta_1)$ is continuous, since f is continuous; also $\Phi(1) = f(P)$, since P has coördinates $\theta_1 = 0$, θ_i arbitrary if $i \geqq 2$. Now it has been proved by Kogbetlianz [18] that the series (5.11) is uniformly summable (C, p) to Φ if $p > \nu$ and the function is continuous. This proves the first assertion of the lemma.

In connection with the second part of the lemma we consider the series

$$(5.13) \qquad \sum_0^\infty (n+\nu) P_n^{(\nu)}(\cos\theta).$$

If $s_n^{(p)}(\cos\theta)$ denotes the Cesàro means of order p of this series we see from (5.5) and (5.9) that

$$(5.14) \qquad H_n^{(p)}(P) = \frac{\Gamma(\nu)}{2\pi^{\nu+1}} \int_{V_k} s_n^{(p)}(\cos\gamma) f(Q)\, dQ.$$

In particular, for $f \equiv 1$ we find

$$(5.15) \qquad 1 = \frac{\Gamma(\nu)}{2\pi^{\nu+1}} \int_{V_k} s_n^{(p)}(\cos\gamma)\, dQ.$$

Now Kogbetlianz has proved [19] that $s_n^{(p)}(\cos\gamma) \geqq 0$ if $p \leqq 2\nu + 1$ and $0 \leqq \gamma \leqq \pi$. The second assertion of the lemma now follows at once from (5.14) and (5.15).

We now turn to the proof of Theorem 10. We begin with the series

$$(5.16) \qquad F(P, \lambda) = \sum_0^\infty e^{-\lambda n(n+2\nu)} a_n(P, f) \qquad \lambda > 0$$

[17] Ervand Kogbetlianz, "Sommabilité des Séries Ultrasphériques," *Journal de Mathématiques*, vol. 3 (1924), p. 110.

[18] Kogbetlianz, *loc. cit.*, p. 118, p. 168.

[19] Kogbetlianz, *loc. cit.*, p. 179.

obtained from (5. 8) by introduction of exponential factors. Concerning this function we make the following statements.

LEMMA 5. *For each P in V_k the function $F(P,\lambda)$ defined by* (5. 16) *is continuous, with continuous derivatives of all orders on the open interval* $0 < 1 < \infty$. *Moreover*

$$\lim \frac{\partial^n F(P,\lambda)}{\partial\lambda^n} = \Delta^n f(P) \quad \text{as } \lambda \to 0 \tag{5.17}$$

and

$$\left| \frac{\partial^n F(P,\lambda)}{\partial\lambda^n} \right| \leqq A \max |\Delta^n f| \qquad 0 < \lambda \leqq 1 \tag{5.18}$$

where A is a constant independent of λ, n, and P.

Granting for a moment the truth of the lemma we can easily deduce from it a proof of Theorem 10. For from the lemma and the conditions of Theorem 10 we conclude by Theorem D that $F(P_0,\lambda) \equiv 0$ if P_0 is in U. Therefore $a_n(P_0,f) = 0$ if P_0 is in U. But $P_n^{(\nu)}(\cos\gamma)$ is a polynomial [20] in $\cos\gamma$, and $\cos\gamma$ is a polynomial in $\sin\theta_i$, $\cos\theta_i$, $\sin\beta_i$, $\cos\beta_i$, $i = 1,\cdots,k$, where θ_i refers to Q and β_i to P in (5. 5). Thus $a_n(P,f)$ is a function of P is analytic on V_k, and since it vanishes on U it vanishes identically. Then by Lemma 4, $f(P) \equiv 0$.

It remains only to prove Lemma 5. The series (5. 16) is a power series in $t = e^{-\lambda}$. Since (5. 8) is summable (C,p) if $p > \nu$ we know that [21]

$$\lim_{n\to\infty} \frac{a_n(P,f)}{n^p} = 0 \qquad p > \nu$$

and from this it follows that (5. 16) is convergent for all positive values of λ (i. e. for $0 \leqq t < 1$). We may obviously differentiate term by term with respect to λ. Doing this we obtain

$$\begin{aligned}\frac{\partial F(P,\lambda)}{\partial\lambda} &= -\lambda \sum_0^\infty n(n+2\nu)e^{-\lambda n(n+2\nu)}a_n(P,f)\\ &= \sum_0^\infty e^{-\lambda n(n+2\nu)}a_n(P,\Delta f)\end{aligned}$$

because of (5. 7). Thus in general

$$\frac{\partial^m F(P,\lambda)}{\partial\lambda^m} = \sum_{n=0}^\infty e^{-\lambda n(n+2\nu)}a_n(P,\Delta^m f), \qquad \lambda > 0. \tag{5.19}$$

---

[20] For the expression of $\cos\gamma$, see Heine, *op. cit.*, p. 458, and the second line on p. 461.

[21] See for example Tomlinson Fort, *Infinite Series*, Oxford (1930), p. 210, Theorem 210.

From this it follows that in the proof of Lemma 5 we may restrict our attention to the case $n = 0$.

If we reorder the series (5.16) according to the rule

$$\sum_0^\infty a_n b_n = \sum_0^\infty (a_0 + \cdots + a_n)(b_n - b_{n+1})$$

applied $p + 1$ times, where we put $b_n = e^{-\lambda n(n+2\nu)}$, we obtain

$$F(P, \lambda) = \sum_{n=0}^\infty (-1)^{p+1} \binom{n+p}{p} H_n^{(p)}(P) D_n^{p+1} b_n \tag{5.20}$$

where $D_n^0 b_n = b_n$, $D_n^1 b_n = b_{n+1} - b_n$, and $D_n^{p+1} b_n = D_n^1(D_n^p b_n)$. The justification of this rearrangement is easily made because of the speed with which b_n and $D_n^{(4)} b_n \to 0$ as $n \to \infty$.

Now we put $p = 2\nu + 1 = k$ and $M = \max | f |$. Then by Lemma 4

$$| F(P, \lambda) | \leqq M \sum_0^\infty \binom{n+k}{k} | D_n^{k+1} b_n$$

and it suffices, to establish (5.18), to show that for some constant A we have

$$K(\lambda) = \sum_0^\infty \binom{n+k}{k} | D_n^{k+1} b_n | \leqq A \qquad 0 < \lambda \leqq 1. \tag{5.21}$$

We write

$$b_n = e^{\lambda\nu^2} e^{-\lambda(n+\nu)^2} = e^{\lambda\nu^2} \phi(n).$$

Then

$$D_n^{k+1} b_n = e^{\lambda\nu^2} \int_0^1 \cdots \int_0^1 \phi^{(k+1)}(n + t_1 + \cdots + t_{k+1}) dt_1 \cdots dt_{k+1}.$$

Since

$$\binom{n+k}{k} \leqq (n+k)^k \leqq (1+k)^k n^k \quad \text{if} \quad n \geqq 1$$

we see that (5.21) will hold if the series

$$\sum_{n=1}^\infty \int_0^1 \cdots \int_0^1 (n + t_1 + \cdots + t_{k+1})^k \times | \phi^{(k+1)}(n + t_1 + \cdots + t_{k+1}) | dt_1 \cdots dt_{k+1}$$

is bounded for $0 < \lambda \leqq 1$. But because of the equation

$$\sum_{n=1}^\infty \int_0^1 \psi(n + t_1) dt_1 = \sum_1^\infty \int_n^{n+1} \psi(s) ds = \int_1^\infty \psi(s) ds$$

the above series is equal to

$$\int_0^1 \cdots \int_0^1 dt_2 \cdots dt_{k+1} \int_1^\infty t_1{}^k \mid \phi^{k+1}(t_1) \mid dt_1 = \int_1^\infty t^k \mid \phi^{k+1}(t) \mid dt.$$

It is therefore sufficient to show that this integral is bounded for $0 < \lambda \leqq 1$. We set $s = \lambda^{1/2}(t + \nu)$ and write

$$\psi(s) = \phi(t) = e^{-s^2}.$$

Then

$$\phi^{(k+1)}(t) = \lambda^{(k+1)/2}\psi^{(k+1)}(s)$$

and therefore, with $s_1 = \lambda^{1/2}(1 + \nu)$,

$$\int_1^\infty t^k \mid \phi^{(k+1)}(t) \mid dt = \int_{s_1}^\infty (s - \nu\lambda^{1/2})^k \mid \psi^{(k+1)}(s) \mid ds.$$

This integral is less than an integral of the form

$$\int_0^\infty Q(\lambda, s) e^{-s^2} ds$$

where $Q(\lambda, x)$ is a polynomial in s and $\lambda^{1/2}$. This is clearly bounded for $0 \leqq \lambda \leqq 1$. Thus (5. 21) and (5. 18) are established.

We have finally to prove (5. 17). Since if $f \equiv 1$ we have $H_n{}^{(k)}(P) \equiv 1$ and $F(P, \lambda) \equiv 1$, we have from (5. 20)

$$1 = \sum_0^\infty (-1)^{k+1} \binom{n+k}{k} D_n{}^{k+1} b_n.$$

Therefore in general

$$F(P, \lambda) - f(P) = \sum_0^\infty (-1)^{k+1} \binom{n+k}{k} \{H_n{}^{(k)}(P) - f(P)\} D_n{}^{k+1} b_n.$$

Now for a fixed P in V_k and a given $\mathcal{E} > 0$ we can choose $N = N(\mathcal{E})$ so that

$$\mid H_n{}^{(k)}(P) - f(P) \mid < \mathcal{E} \quad \text{if} \quad n \geqq N.$$

Since $b_n = e^{-\lambda n(n+2\nu)}$ it is clear that $D_n{}^{k+1} b_n \to 0$ as $\lambda \to 0$ for a fixed n. Therefore, from (5. 21) it follows that

$$\overline{\lim} \mid F(P, \lambda) - f(P) \mid \leqq \mathcal{E} A \quad \text{as} \quad \lambda \to 0.$$

This proves the validity of (5. 17), since $\mathcal{E}$ was arbitrary. Thus theorem 10 is completely proved.

PRINCETON UNIVERSITY.

PROBLEMS OF THE CALCULUS OF VARIATIONS WITH PRESCRIBED TRANSVERSALITY CONDITIONS.*

By Julian D. Mancill.

Problems of the calculus of variations in $(x, y_1, y_2, \cdots, y_n)$-space for which a prescribed relation exists between the directions of the extremals and the transversal directions were first studied by Rawles.[1] More recently, La Paz,[2] using a method and point of view quite different from that of Rawles, has given a rather complete treatment of the problem in non-parametric form.

In the present paper we use a method similar to that of La Paz but avoid his very intricate treatment of an associated system of non-homogeneous partial differential equations by reducing the problem to a very simple total differential equation. The method applies with equal facility to parametric and non-parametric problems of the calculus of variations in space of any number of dimensions. We shall consider the problem in parametric form and derive necessary and sufficient conditions in order that a transversality relation belong to a problem of the calculus of variations. Finally, we shall obtain the most general integrand function of a problem of the calculus of variations to which a given transversality relation belongs.

We shall consider the integral

$$I = \int_{t_1}^{t_2} F(x, x')\,dt, \tag{1}$$

where the integrand function $F(x_1, \cdots, x_n, x'_1, \cdots, x'_n) \equiv F(x, x')$ satisfies the usual continuity and homogeneity properties in a fundamental region R'.[3] For a problem of minimizing the integral (1), an extremal E through the point (x) in the direction (x') is cut transversally by the hyper-plane of directions $(\bar{x}')$ defined through this point by the equation

* Received May 11, 1938.

[1] *Transactions of the American Mathematical Society*, vol. 30 (1928), pp. 765-784. For the earlier writers on this problem in the plane case, see Stromquest, *Transactions of the American Mathematical Society*, vol. 7 (1906), p. 181; and Bliss, *Annals of Mathematics* (2), vol. 9 (1907), p. 134.

[2] *Bulletin of the American Mathematical Society*, vol. 36 (1930), pp. 674-680.

[3] In this connection, see Bolza, *Lectures on the Calculus of Variations*, University of Chicago, pp. 17-21. In this paper we shall assume that the function F is of class C''' in a region R and all $(x') \neq (0)$. For a definition of the term class as here used see Bolza, loc. cit., p. 116.

$$\bar{x}'_\mu F_{x'_\mu}(x, x') = 0. \tag{2}$$

Here, as elsewhere in this paper, a repeated Greek letter is an umbral index indicating a summation with the range 1 to n unless otherwise specified. This definition of transversality simply defines what is meant by transversality for every element (x, x') in the fundamental region R' of the integral (1).

We shall restrict our attention to that portion of the region R' *in which the function* $F(x, x')$ *is different from zero.* In this subregion of R' the functions

$$t_i = F_{x'_i}(x, x') \qquad (i = 1, 2, \cdots, n) \tag{3}$$

are of class C'' and not all zero for any element (x, x'), as follows from the homogeneity property

$$x'_\mu F_{x'_\mu}(x, x') = F(x, x'). \tag{4}$$

Hence, in the subregion specified, we have

$$x'_\mu F_{x'_\mu}(x, x') \neq 0. \tag{5}$$

We shall say that the functions (3) *define* the transversality relation (2) and that this transversality *belongs* to the calculus of variations problem (1).

The condition (2) establishes a certain relation between the element (x, x') through the point (x) and the hyper-plane of directions $(\bar{x}')$ whose normal has the direction (t) through the point. If n given functions $T_i(x, x')$, $i = 1, 2, \cdots, n$, of class C'' in a region S' are to define a transversality

$$\bar{x}'_\mu T_\mu(x, x') = 0 \tag{6}$$

belonging to a problem (1) with fundamental region S' in which the integrand function is not zero, then the equation (6) and the equation (2) define the same hyper-plane of directions $(\bar{x}')$ for every element (x, x') in S'. That is, there must exist functions $F(x, x')$ and $K(x, x')$ different from zero in S' and satisfying the system of equations

$$F_{x'_i}(x, x') = K(x, x') T_i(x, x') \qquad (i = 1, 2, \cdots, n) \tag{7}$$

and the homogeneity condition (4). It is easily verified that in order that the homogeneity property (4) be satisfied by the function $F(x, x')$, we must have

$$K(x, x') = F(x, x') / x'_\mu T_\mu(x, x'), \tag{8}$$

since $x'_\mu T_\mu \neq 0$ as follows from the relations (4) and (7). Thus it follows from (7) that

$$F_{x'_i}(x, x') = F(x, x') T_i(x, x') / x'_\mu T_\mu(x, x') \qquad (i = 1, 2, \cdots, n). \tag{9}$$

If these partial derivatives be substituted in

$$dF = F_{x'_\mu}(x, x')\,dx'_\mu$$

where F is regarded as a function of the n variables (x') and the n parameters (x), we obtain the total differential equation

$$dF/F = T_\mu dx'_\mu / x'_\mu T_\mu. \tag{10}$$

The left member of this equation is the exact differential of the function $\log F$. It is found that necessary conditions for the right member to be the differential of a function $H(x, x')$ considered as a function of the variables (x'), are that the functions T_i satisfy the $n(n-1)/2$ relations

$$R_{m,k} = x'_\mu[T_m \partial T_\mu/\partial x'_k - T_k \partial T_\mu/\partial x'_m + T_\mu(\partial T_k/\partial x'_m - \partial T_m/\partial x'_k)] = 0,$$

where $m, k = 1, 2, \cdots, n$, $m < k$. These conditions are also sufficient if suitable connectivity properties are assumed for the region S'. An easy calculation shows that the expressions $R_{m,k}$ can be expressed as follows:

$$\begin{aligned} R_{m,k} = {} & T_m(x'_\mu \partial T_k/\partial x'_\mu) - T_k(x'_\mu \partial T_m/\partial x'_\mu) \\ & + x'_\nu[T_m(\partial T_\nu/\partial x'_k - \partial T_k/\partial x'_\nu) + T_k(\partial T_m/\partial x'_\nu - \partial T_\nu/\partial x'_m) \\ & + T_\nu(\partial T_k/\partial x'_m - \partial T_m/\partial x'_k)] = 0, \end{aligned} \tag{11}$$

where ν is an umbral index with range 1 to n excluding m and k.

If the conditions (11) are satisfied in S' then there exists a function

$$F(x, x') = G(x)e^{H(x,x')} \tag{12}$$

satisfying the equations (7) with $K(x, x')$ defined by equation (8). Then the functions T_i satisfy the $n(n-1)(n-2)/6$ equations

$$\begin{aligned} T_r(\partial T_t/\partial x'_s - \partial T_s(\partial x'_t) + T_s(\partial T_r/\partial x'_t - \partial T_t/\partial x'_r) \\ + T_t(\partial T_s/\partial x'_r - \partial T_r/\partial x'_s) = 0, \end{aligned} \tag{13}$$

where the suffixes r, s, and t represent all possible combinations of three suffixes from among $1, 2, \cdots, n$.[4] It can be shown that only $(n-1)(n-2)/2$ of these conditions are independent, one set of which may be obtained by fixing one of the suffixes and letting the other two represent all possible combinations of two suffixes from among the remaining $n-1$ numbers. Therefore, if the equations (11) are satisfied so are the equations

$$T_m(x'_\mu \partial T_k/\partial x'_\mu) - T_k(x'_\mu \partial T_m/\partial x'_\mu) = 0, \tag{14}$$

---

[4] Forsyth, *Differential Equations*, Macmillan and Co. (1888), pp. 261-262.

where $m, k = 1, 2, \cdots, n$, $m < k$. Conversely, if the equations (13) and (14) are satisfied so are equations (11). Consequently, the two sets of conditions (13) and (14) are equivalent to the set (11).

Now, the set of equations (14) will be satisfied if and only if the n functions T_i are such that

$$x'_\mu \partial T_i/\partial x'_\mu = h(x, x')T_i \qquad (i = 1, 2, \cdots, n),$$

that is, if and only if the functions T_i considered as functions of the n variables (x') are solutions of the same partial differential equation

$$x'_\mu \partial T/\partial x'_\mu = h(x, x')T. \tag{15}$$

To obtain the most general integral of this linear equation we write the subsidiary equations

$$dx'_1/x'_1 = dx'_2/x'_2 = \cdots = dx'_n/x'_n = dT/h(x, x')T$$

and obtain n independent integrals of these, namely

$$x'_1/x'_n = a_1,\ x'_2/x'_n = a_2, \cdots,\ x'_{n-1}/x'_n = a_{n-1},\ T/U(x, x') = a_n,$$

where the a_i are constant with respect to the variables (x'). The last integral is obtained by making the substitution $x'_i = a_i x'_n$, $i = 1, 2, \cdots, n-1$, in $h(x, x')$ and integrating the equation

$$h\,dx'_n/x'_n = dT/T$$

and then replacing the a_i by their values x'_i/x'_n, $i = 1, 2, \cdots, n-1$. Then the most general solution of the equation (15) is of the form

$$T = U(x, x')W(x_1, \cdots, x_n, u_1, \cdots, u_{n-1}), \quad u_k = x'_k/x'_n,$$
$$(k = 1, 2, \cdots, n-1),$$

where W is an arbitrary function of its arguments. Conversely, the function

$$T = R(x, x')W(x_1, \cdots, x_n, u_1, \cdots, u_{n-1})$$

where R and W are arbitrary functions of their arguments is a solution of the equation (15) for $h = x'_\mu R_{x'_\mu}/R$. Therefore, the conditions (14) will be satisfied if and only if the functions T_i are of the form

$$T_i = R(x, x')W_i(x_1, \cdots, x_n, u_1, \cdots, u_{n-1}), \quad u_k = x'_k/x'_n,$$
$$(k = 1, 2, \cdots, n-1;\ i = 1, 2, \cdots, n), \tag{16}$$

where R and W_i are arbitrary functions of their arguments.

We may summarize our results in the following theorem:

THEOREM. *If $T_i(x, x')$, $i = 1, 2, \cdots, n$, are of class C'' in a region S' of (x, x')-values, then necessary and (if S' possesses suitably simple connectivity properties) sufficient conditions that these functions define a transversality relation* (2) *for a problem* (1) *whose integrand F is different from zero are that these functions satisfy the inequality $x'_\mu T_\mu(x, x') \neq 0$, be of the form* (16), *and satisfy $(n-1)(n-2)/2$ independent relations of the form* (13). *Whenever these conditions are satisfied, the corresponding integrand function is of class C''' and of general form* (12), *in which $H(x, x')$ is defined by the line integral*

$$H(x, x') = \int_{x'_0}^{x'} T_\mu dx'_\mu / x'_\mu T_\mu,$$

where (x'_0) is any fixed admissible element of S', and $G(x)$ is different from zero and of class C''' but is otherwise an arbitrary function of its arguments.

The fact that the integrand function F as defined in the Theorem satisfies the relation

$$F(x, kx') = kF(x, x') \qquad (k > 0)$$

is easily seen by making the substitution $x'_i = kx'_i$, $i = 1, 2, \cdots, n$, in (10) and making use of the property (16) of the functions T_i.

As an illustration of the Theorem let us obtain the most general problem (1) for which transversality is equivalent to orthogonality.[5] In this case $T_i = x'_i$, $i = 1, 2, \cdots, n$, which satisfy all the conditions of the Theorem for all $(x') \neq (0)$. The function $H(x, x')$ takes the form

$$H(x, x') = (1/2) \log (x'_\mu x'_\mu),$$

if the lower limit in the integral is properly chosen. Therefore, the integrand function $F(x, x')$ is of the form

$$F(x, x') = G(x)(x'_\mu x'_\mu)^{1/2}.$$

UNIVERSITY OF ALABAMA.

[5] Mancill, *Bulletin of the American Mathematical Society*, vol. 43 (1937), p. 30, Abstract 41.

NORMAL MATRICES OVER AN ARBITRARY FIELD OF CHARACTERISTIC ZERO.*

By John Williamson.

Introduction. A matrix A with elements in the complex number field is said to be a normal matrix, if $AA^* = A^*A$, where A^* is the conjugate transposed of the matrix A. A necessary and sufficient condition that A be a normal matrix is that there exist a unitary matrix U, such that $U^*AU = D$, where D is a diagonal matrix. Further a matrix A is normal, if, and only if, $A^* = f(A)$, where $f(x)$ is a polynomial in x.[1] From this it is possible to make a satisfactory definition of normality with respect to any non-singular hermitian matrix H. A matrix A is said to be normal with respect to the hermitian matrix H, if $AH = Hf(A^*)$. The simplicity of the canonical form of a normal matrix A under unitary transformation suggests the following problems. What are the possible canonical forms for a matrix A, normal with respect to H, under similarity transformations by matrices which are conjunctive automorphs of H? What are necessary and sufficient conditions that two matrices, both normal with respect to H, be similar under transformations by matrices, which are conjunctive automorphs of H. These problems were discussed in a previous paper.[1] It is our intention here to consider the corresponding problems where the matrices under consideration are matrices over a field of characteristic zero.

1. Definitions. Let K be a field of characteristic zero, over which is defined an automorphism of period one or two. If under this automorphism an element a of K corresponds to an element $\bar{a}$ of K, in both cases $\bar{\bar{a}} = a$. If $A = (a_{ij})$, $i = 1, 2, \cdots, m$; $j = 1, 2, \cdots, n$, is a matrix with elements in K, the matrix A^*, the conjugate transposed of A, is the matrix whose element in the i-th row and j-th column is $\bar{a}_{ji}$. If the automorphism over K is of period one, A^* is simply A' the transposed of A. The matrix A is said to be hermitian if $A^* = A$ and anti-hermitian if $A^* = -A$. If the automorphism over K is of period one, an hermitian matrix is symmetric and an anti-hermitian one, skew symmetric. In the sequel all matrices are matrices over the field K.

---

* Received July 30, 1938.

[1] John Williamson, "Matrices normal with respect to an hermitian matrix," *American Journal of Mathematics*, vol. 60 (April, 1938), pp. 355-374.

DEFINITION. *If H is a non-singular hermitian or anti-hermitian matrix, so that $H^* = \epsilon H$, $\epsilon = \pm 1$, and $AH = Hf(A^*)$, where $f(x)$ is a polynomial in the ring $K[x]$, the matrix A is normal with respect to H.*

Let A be normal with respect to H, so that

$$(1) \qquad AH = Hf(A^*).$$

Then, since $H^* = \epsilon H$,

$$HA^* = \bar{f}(A)H = H\bar{f}\{f(A^*)\} \text{ by } (1),$$

and, since H is non-singular,

$$(2) \qquad A^* = \bar{f}\{f(A^*)\} \text{ and } A = f\{\bar{f}(A)\}.$$

On writing $B = f(A^*)$, we have as a consequence of (1) and (2)

$$(3) \quad AH = HB,\ B^*H = HA^*,\ B^* = \bar{f}(A),\ A^* = \bar{f}(B),\ A = f(B^*).$$

It is easily shown that, if

$$PAP^{-1} = A_1 \text{ and } PHP^* = H_1,\ A_1H_1 = H_1f(A^*),$$

so that A_1 is normal with respect to H_1.

DEFINITION. *Two matrices A_1 and A_2 are H-equivalent, if there exists a non-singular matrix P such that $PHP^* = H$ and $PA_1P^{-1} = A_2$.*

Consequently we have

LEMMA 1. *If A_1 and A_2 are H-equivalent and A_1 is normal with respect to H, then A_2 is normal with respect to H and, if $A_1H = Hf(A_1^*)$, $A_2H = Hf(A^*_2)$.*

2. Statement of problem. Let A_1 and A_2 be two similar matrices. Then A_1 and A_2 are both similar to the same matrix Q and there exist two non-singular matrices P_1 and P_2, such that

$$(4) \qquad A_i = P_iQP_i^{-1}, \qquad (i = 1, 2).$$

If A_1 and A_2 are both normal with respect to H, by Lemma 1, Q is normal with respect to both of the matrices

$$(5) \qquad S_i = P_iHP^*_i, \qquad (i = 1, 2).$$

The matrices S_i in (5) are both hermitian, when H is hermitian, and anti-

hermitian, when H is anti-hermitian. The proof of the following fundamental theorem is omitted.[2]

THEOREM 1. *Necessary and sufficient conditions that two matrices, both normal with respect to H, be H-equivalent is that there exist a non-singular matrix C such that*

$$CQ = QC \text{ and } CS_1C^* = S_2,$$

where Q, S_1 and S_2 are any matrices, which satisfy (5) *and* (6).

The matrix Q is any matrix similar to A_1 and A_2 and is therefore at our disposal. In order to determine when A_1 and A_2 are H-equivalent we need only determine when S_1 is conjunctively equivalent to S_2 under transformations by matrices commutative with Q. As is usual in such cases, we may treat A_1 and A_2 separately and reduce S_1 (and S_2) to a canonical form under conjunctive transformations by matrices commutative with Q. Accordingly we drop the suffix one and write A, S for A_1, S_1 respectively. The matrices Q, S thus obtained will be called a *canonical form* for A, H.

3. Reduction to special cases. The matrix Q is now any matrix similar to A and therefore the invariant factors of $Q - xE$ are the same as those of $A - xE$. The powers of the distinct irreducible polynomials, which are divisors of the invariant factors of $A - xE$, will be called the *elementary factors* of $A - xE$ or for brevity the elementary factors of A. Therefore the elementary factors of Q are the same as the elementary factors of A. Further, if $[p(x)]^e$ is an elementary factor of A, we shall call $p(x)$ a *characteristic factor* of A.

We first choose Q to be the *diagonal block matrix*,

$$Q = [Q_1, Q_2, \cdots, Q_k],$$

where each characteristic factor of Q_i has the same value $p_i(x)$ and $p_i(x) \neq p_j(x)$, when $i \neq j$. Then, if $PAP^{-1} = Q$ and $PHP^* = S$, by Lemma 1,

$$QS = Sf(Q^*) = SM, \tag{6}$$

where M is the diagonal block matrix

$$M = [M_1, M_2, \cdots, M_k]$$

and

$$M_j = f(Q^*_j).$$

---

[2] See John Williamson; *loc. cit.*, p. 260.

If the elementary factors of Q_i are $p_i(x)^{e_{ij}}$, $j=1,2,\cdots,r_i$, we may suppose Q_i to be in the Wedderburn canonical form,[3] so that

(7) $$Q_i=[N_{i1},N_{i2},\cdots,N_{i,t_i}],$$

where

(8) $$N_{ij}=p_iE_{ij}+U_{ij}.$$

In (8) p_i is the companion matrix of $p_i(x)$, p_iE_{ij} is the direct product of p_i with the unit matrix of order e_{ij} and U_{ij} the direct product of the unit matrix of the same order as p_i with the auxiliary unit matrix of order e_{ij}. Then

$$M_i=f(Q^*_i)=[f(N^*_{ij})]=[R_{ij}E_{ij}+V_{ij}],\qquad (j=1,2,\cdots,t_i),$$

where V_{ij} is nilpotent. If the characteristic equation of R_{ij} is irreducible, the number of elementary factors of M_i is the same as that of Q_i; otherwise the number of elementary factors of M_i is more than that of Q_i. Since M is similar to Q, the number of elementary factors of M is the same as that of Q and therefore the characteristic equation of R_{ij} is irreducible. Hence the characteristic factors of M_i all have the same value $p_j(x)$. Two distinct cases arise:

Case (i). The characteristic factors of M_i all have the same value $p_j(x)$ and $i\neq j$.

Case (ii). The characteristic factors of M_i all have the same value $p_i(x)$.

Let

$$S=(S_{rs}),\qquad (r,s=1,2,\cdots,k),$$

be a *partition* of S *similar* to that of Q; i.e. S_{rs} is a matrix with the same number of rows as Q_r and the same number of columns as Q_s. Then, as a consequence of (6),

(9) $$Q_rS_{rs}=S_{rs}M_s,\qquad (r,s=1,2,\cdots,k).$$

Case (i). Since M_i has no characteristic factor in common with Q_s when $s\neq j$, $S_{si}=0$ when $s\neq j$. Further, since $S=\epsilon S^*$,

$$S_{is}=\epsilon S^*_{si}=0,$$

when $s\neq j$.

Finally, since S is non-singular, S_{ji} must be square and non-singular and so must S_{ij}. Equation (9) with $r=i$ and $s=j$, shows that M_j is similar to Q_i and therefore that all characteristic factors of M_j have the same value $p_i(x)$.

---

[3] J. M. Wedderburn, "Lectures on matrices," *Colloquium Publications* (1934), pp. 123-124.

Hence in case (i), Q_i and Q_j are of the same order and the exponents of the elementary factors of Q_i are the same as those of Q_j.

Case (ii). As a consequence of (9), $S_{is} = 0$, unless $s = i$. Since S is non-singular, S_{ii} is non-singular.

Accordingly after a rearrangement of the rows of Q, M, and S and the same rearrangement of the columns, we see that Q, M and S are similarly partitioned diagonal block matrices. The blocks are of two distinct types:

Type (i).

$$[Q_i, Q_j],\ [M_i, M_j],\ \begin{pmatrix} 0 & S_{ij} \\ S_{ji} & 0 \end{pmatrix},$$

where

$$\begin{pmatrix} Q_i & 0 \\ 0 & Q_j \end{pmatrix}\begin{pmatrix} 0 & S_{ij} \\ S_{ji} & 0 \end{pmatrix} = \begin{pmatrix} 0 & S_{ij} \\ S_{ji} & 0 \end{pmatrix}\begin{pmatrix} M_i & 0 \\ 0 & M_j \end{pmatrix};$$

Type (ii). Q_i, M_i, S_{ii}, where $Q_i S_{ii} = S_{ii} M_i$.

Since any matrix commutative with Q is also a diagonal block matrix partitioned similarly to Q,[4] we may treat each block separately. We first consider those of type (i).

4. Type (i). Let E be the identity matrix of the same order as Q_i and let $R = \begin{pmatrix} E & 0 \\ 0 & S_{ji}^{-1} \end{pmatrix}$. Then

$$R[Q_i, Q_j]R^{-1} = [Q_i, S_{ji}^{-1} Q_j S_{ji}] = [Q_i, M_i] = [Q_i, f(Q^*_i)], \tag{10}$$

and

$$R\begin{pmatrix} 0 & S_{ij} \\ S_{ji} & 0 \end{pmatrix} R^* = \begin{pmatrix} E & 0 \\ 0 & S_{ji}^{-1} \end{pmatrix}\begin{pmatrix} 0 & S_{ij} \\ S_{ji} & 0 \end{pmatrix}\begin{pmatrix} E & 0 \\ 0 & \epsilon S_{ij}^{-1} \end{pmatrix} = \begin{pmatrix} 0 & \epsilon E \\ E & 0 \end{pmatrix}. \tag{11}$$

Since, by (10), $[Q_i, Q_j]$ is similar to $[Q_i, f(Q^*_i)]$, we may replace the block $[Q_i, Q_j]$ in Q by $[Q_i, f(Q^*_i)]$. If this is done, the corresponding block in S is, as a consequence of (11), $\begin{pmatrix} 0 & \epsilon E \\ E & 0 \end{pmatrix}$. We have therefore

Result (a). *Let $p_i(x)$ and $p_j(x)$ be two distinct characteristic factors of the matrix A, which is normal with respect to H. If p_i and p_j are the companion matrices of $p_i(x)$ and $p_j(x)$ respectively and if $f(p^*_j)$ satisfies $p_i(x) = 0$, then $f(p^*_i)$ satisfies $p_j(x) = 0$. Corresponding to these two characteristic factors there is in the canonical form of Q a block $[Q_i, f(Q^*_i)]$ and in S the block $\begin{pmatrix} 0 & \epsilon E \\ E & 0 \end{pmatrix}$.*

---

[4] John Williamson, "The idempotent and nilpotent elements of a matrix," *American Journal of Mathematics*, vol. 58 (October, 1936), no. 4, p. 757.

The matrix Q_i in the above is not unique but may be replaced by any matrix similar to it. If, however, Q_i is taken in the Wedderburn canonical form, for instance, the canonical form above is uniquely determined. If $p_i(x)$ is a characteristic factor of A and if $f(p^*_i)$ does not satisfy $p_i(x) = 0$, then for some value of j, distinct from i, $f(p^*_j)$ satisfies $p_i(x) = 0$ and $f(p^*_i)$ satisfies $p_j(x) = 0$. Therefore we have

THEOREM 2. *Let A be normal with respect to H, so that $AH = Hf(A^*)$. Let the characteristic factors of A be $p_i(x)$, $i = 1, 2, \cdots, k$ and let p_i be the companion matrix of $p_i(x)$. If for no value of i, $i = 1, 2, \cdots, k$, $f(p^*_i)$ satisfies the equation $p_i(x) = 0$, there exists a non-singular matrix R such that*

$$RAR^{-1} = \begin{pmatrix} F & 0 \\ 0 & f(F^*) \end{pmatrix} \text{ and } RHR^* = \begin{pmatrix} 0 & \epsilon E \\ E & 0 \end{pmatrix}. \tag{12}$$

The matrix F in (12) is of order one-half that of A while E is the unit matrix of the same order as F. The matrix F may be replaced by any matrix similar to it and, if F is chosen in the Wedderburn canonical form, then F is unique. Accordingly (12) gives a canonical form for the matrices A and H when all characteristic factors of A are of the type (i). The canonical form (12) is determined completely by the invariant factors of A and the polynomial $f(x)$.

We have as an immediate corollary,

COROLLARY 1. *If two similar matrices, which are both normal with respect to the same hermitian or anti-hermitian matrix H, have all their characteristic factors of type* (i), *they are also H-equivalent.*

5. Proof of lemmas. Before proceeding to the consideration of the case, in which the characteristic factors of A are of type (ii), we prove some lemmas.

LEMMA 11. *Let $Q = [Q_1, Q_2]$, $M = [M_1, M_2]$ and $S = (S_{ij})$, $i, j = 1, 2$, be similar partitions of the matrices Q, M and S and let $QS = SM$. If $S^* = \epsilon S$ and S_{11} is non-singular, there exists a non-singular matrix R, such that $RQ = QR$ and $RSR^* = [S_{11}, T_{22}]$.*

Since

$$Q_iS_{ij} = S_{ij}M_j, \qquad (i, j = 1, 2),$$
$$S_{21}S_{11}^{-1}Q_1 = S_{21}M_1S_{11}^{-1} = Q_2S_{21}S^{-1}.$$

Therefore, if E_i is the unit matrix of the same order as Q_i and

$$R = \begin{pmatrix} E_1 & 0 \\ -S_{21}S_{11}^{-1} & E_2 \end{pmatrix},$$

$RQ = QR$. Further

$$R^* = \begin{pmatrix} E_1 & -(S_{21}S_{11}^{-1})^* \\ 0 & E_2 \end{pmatrix} = \begin{pmatrix} E_1 & -\epsilon^2 S_{11}^{-1}S_{12} \\ 0 & E_2 \end{pmatrix} = \begin{pmatrix} E_1 & -S_{11}^{-1}S_{12} \\ 0 & E_2 \end{pmatrix}.$$

Consequently

$$RSR^* = [S_{11}, S_{22} - S_{21}S_{11}^{-1}S_{12}]$$

and the lemma is proved.

Let U be the auxiliary unit matrix of order n and V the auxiliary unit matrix of order m. Let $p(x)$ be an irreducible polynomial of the ring $K[x]$ and p be the companion matrix of $p(x)$. If θ is a zero of $p(x)$, then the totality of matrix polynomials in p with coefficients in K is a field simply isomorphic with the simple algebraic adjunction field $K(\theta)$. If $a_i = a_i(p)$ is a polynomial in p, by a_iU we mean the matrix obtained from $a_i(\theta)U$ by replacing θ by the matrix p. In other words, a_iU is the direct product of a_i by U written in a definite way.

We now consider the matrix equation

$$\phi(U)D = D\psi(V), \tag{13}$$

where

$$D = (d_{ij}), \qquad (i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, m),$$

and

$$\phi(U) = pE_n + \sum_{i=1}^{n-1} b_iU^i, \qquad \psi(V) = pE_m + \sum_{j=1}^{m-1} c_jV^j,$$

$b_i = b_i(p)$, $c_j = c_j(p)$, $b_1c_1 \neq 0$ and E_k is the unit matrix of order k. Then each element d_{ij} is a polynomial in p and we have

LEMMA 111. *If D satisfies* (13), *the first column of D is zero, when $m > n$, and the last row is zero, when $m < n$. If $m = n$, D is non-singular if, and only if, $d_{11} \neq 0$.*[5]

Further, if $m = n$, so that D is a square matrix,

$$D = D_0 + D_1 + \cdots + D_{n-1},$$

where the only non-zero elements of D_j are those d_{rs} for which $s - r = j$. It now follows from (13) that

$$\sum_{i=1}^{n-1} b_iU^i \sum_{j=0}^{n-1} D_j = \sum_{j=0}^{n-1} D_j \sum_{i=1}^{n-1} c_iU^i$$

and, from the nature of the matrices D_j, that

[5] For proof see (1) page 363. The proof is the same as that of lemma 4 except that b_i and c_j are now polynomials in p.

$$(14)\qquad \sum_{a=1}^{s} b_a U^a D_{s-a} = \sum_{a=1}^{s} c_a D_{s-a} U^a, \qquad (s = 1, 2, \cdots, n-1).$$

Let $D_1 = D_2 = \cdots = D_{r-1} = 0$. Then from the first r of equations (14) we have

$$(15)\qquad b_k U^k D_0 = c_k D_0 U^k, \qquad (k = 1, 2, \cdots, r).$$

If now

$$b_{r+2} = b_{r+3} = \cdots = b_{n-1} = c_{r+1} = c_{r+2} = \cdots = c_{n-1} = 0,$$

the $(r+1)$-th equation in (14) becomes

$$(16)\qquad b_1 U D_r + b_{r+1} U^{r+1} D_0 = c_1 D_r U,$$

while the remaining equations take the form

$$\sum_{a=1}^{r+1} b_a U^a D_{s-a} = \sum_{a=1}^{r} c_a D_{s-a} U^a,$$

or

$$(17)\qquad b_1 U D_{s-1} - c_1 D_{s-1} U = \sum_{a=2}^{r} c_a D_{s-a} U^a - \sum_{a=2}^{r+1} b_a U^a D_{s-a}.$$

If D_j is known for $j < s-1$, (17) may be solved for D_{s-1}. In fact if d_i denotes the element, which is different from zero, in the i-th row of D_{s-1}, the non-zero element in the i-th row of the matrix on the left of (17) is $b_1 d_{i+1} - c_1 d_i$. Since $b_1 \neq 0$, d_2, d_3 etc. can be determined successively in terms of d_1 and the elements of D_j, $j < s-1$. We have therefore proved

LEMMA IV. *If* (15) *and* (16) *are satisfied, and D_0 is non-singular, there exists a matrix*

$$D = D_0 + D_r + \cdots + D_{n-1}$$

such that

$$D \sum_{i=1}^{r} c_i U^i D^{-1} = \sum_{j=1}^{r+1} b_j U^j.$$

6. Case (ii). We now consider the case in which Q_i is of type (ii) so that $Q_i S_{ii} = S_{ii} M_i$. For convenience we temporarily drop all suffixes i so that $QS = SM$. Let the elementary factors of Q be $[p(x)]^{e_i}$, $i = 1, 2, \cdots, k$ where $e_1 \geq e_2 \geq \cdots \geq e_k$. Let E_i be the unit matrix of order e_i and U_i the auxiliary unit matrix of the same order. Then, if p is the companion matrix of $p(x)$ and $Q_i = pE_i + U_i$, $i = 1, 2, \cdots, k$, $Q = [Q_1, Q_2, \cdots, Q_k]$. Moreover,

$$M = f(Q^*) = [M_1, M_2, \cdots, M_k],$$

where

$$M_i = f(Q^*_i) = f(p^*)E_i + f'(p^*)U'_i + \cdots + 1/(e_i - 1)!\, f^{(e_i-1)}(p^*)U'^{e_i-1}_i$$

$$(18) \qquad = qE_i + \sum_{j=1}^{e_i-1} a_j U'^j_i.$$

The matrices a_j in (18) are completely determined by the matrix p^* and the polynomial $f(x)$ and are therefore independent of i. Since Q is similar to M and M has the single characteristic factor $p(x)$, the minimal equation of q is $p(x) = 0$. Hence p is similar to q and there exists a non-singular matrix y such that

$$(19) \qquad y^{-1}py = q.$$

As a consequence of (3), $p^* = \bar{f}(q)$, so that the a_j in (18) are also polynomials in q and may be written

$$(20) \qquad a_j = a_j(q).$$

Therefore by (19) and (20)

$$(21) \qquad ya_j(q) = a_j(p)y.$$

Let T_i be the counter unit matrix of order e_i so that

$$(22) \qquad T_i^2 = E_i, \qquad T_iU'_i = U_iT_i.$$

If

$$\begin{aligned} T &= [T_1, T_2, \cdots, T_k] \quad \text{and} \quad D = Sy^{-1}T, \\ QDyT &= DyTM \\ &= D[yT_iM_i] = D[N_iyT_i] = D[N_i]yT, \ (i = 1, 2, \cdots, k), \\ &= DNyT, \end{aligned}$$

where, as a consequence of (18), (21) and (22),

$$N_i = pE_i + \sum_{j=1}^{e_i-1} a_j(p)U_i^j.$$

Therefore

$$QD = DN.$$

If

$$D = (D_{ij}), \qquad (i, j = 1, 2, \cdots, k),$$

is a partition of D similar to that of Q, we have

$$(23) \qquad Q_iD_{ij} = D_{ij}N_j, \qquad (i, j = 1, 2, \cdots, k).$$

Each of the equations (23) is of the same type as (13) and therefore the form of D_{ij} is known. Let $e_1 = e_2 = \cdots = e_c > e_{c+1}$ and let d_{ij} denote the

element in the first row and the first column of D_{ij}. If $d_{11} \neq 0$, by Lemma III, D_{11} is non-singular. If $d_{11} = 0$ and $d_{jj} \neq 0$, $j \leq c$, by an interchange of rows and the same interchange of columns we may replace D_{11} by D_{jj} without disturbing Q or N, so that again we may suppose D_{11} to be non-singular. Finally let $d_{11} = d_{22} = \cdots = d_{cc} = 0$. Since by Lemma III the first column of D_{j1} is zero, when $j > c$, and D is non-singular, for at least one value of j, $1 < j \leq c$, $d_{j1} \neq 0$. We may suppose without any loss of generality that $d_{21} \neq 0$. If $\eta(p)$ is any polynomial in p, the matrix

$$R = \left[\begin{pmatrix} E_1 & \eta(p)E_1 \\ 0 & E_1 \end{pmatrix}, I\right],$$

where I is the unit matrix of order $e_3 + e_4 + \cdots + e_k$, is commutative with Q.

If $RDyTR^* = WyT$,

$$\begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}\begin{pmatrix} yT_1 & 0 \\ 0 & yT_1 \end{pmatrix} = \begin{pmatrix} E_1 & \eta(p)E_1 \\ 0 & E_1 \end{pmatrix}\begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}\begin{pmatrix} yT_1 & 0 \\ 0 & yT_1 \end{pmatrix}\begin{pmatrix} E_1 & 0 \\ \bar{\eta}(p^*)E_1 & E_1 \end{pmatrix},$$

so that

$$W_{11} = D_{11} + \eta(p)D_{21} + D_{12}y\bar{\eta}(p^*)y^{-1} + D_{22}\eta(p)y\bar{\eta}(p^*)y^{-1}. \tag{24}$$

But $p^*y = f(q)y = yf(p) = yq^*$ by (3) and (19), so that (24) becomes

$$W_{11} = D_{11} + \eta(p)D_{21} + D_{12}\bar{\eta}(q^*) + D_{22}\eta(p)\bar{\eta}(q^*).$$

Finally, if w_{11} is the element in the first row and column of W_{11} we have, since $d_{11} = d_{22} = 0$,

$$w_{11} = \eta(p)d_{21} + d_{12}\bar{\eta}(q^*). \tag{25}$$

Since R is commutative with Q we may replace D by W. Accordingly by Lemma III, if $w_{11} \neq 0$, W_{11} is non-singular and we may therefore suppose that D_{11} is non-singular. There remains the possibility that w_{11} is zero for every choice of $\eta(p)$. This cannot be the case, if the automorphism over the field K is of period two. For otherwise we would have in particular

$$\theta d_{21} + \bar{\theta} d_{12} = 0 \tag{26}$$

for every scalar matrix θ, which, when $\theta = 1$, yields $d_{21} = -d_{12}$. Since $d_{21} \neq 0$, (26) reduces to $\theta = \bar{\theta}$ for every θ in K and this is a contradiction of the fact that the automorphism over K is actually of period two. If the automorphism over K is of period one and (25) is zero for every $\eta(p)$, then $\eta(p) = \eta(q')$ for every polynomial $\eta(p)$ and therefore $p = q'$ and $q = p'$. Consequently we have proved

LEMMA V. *The matrix D_{11} may be taken as non-singular unless the automorphism over the field K is of period one, $d_{12} = -d_{21}$ and $q = p'$.*

Since D_{12} satisfies (23), it follows from (14) that the principal diagonal of D_{12} is $[d_{12}, ad_{12}, \cdots, a^{n-1}d_{12}]$ and that of D_{21} is $[d_{21}, ad_{21}, \cdots, a^{n-1}d_{21}]$, where for simplicity of notation we have written

$$a = a_1(p) \text{ and } n = e_1.$$

Since $S^* = \epsilon S$, $(DyT)^* = \epsilon(DyT)$ and

$$(D_{21}yT_1)^* = \epsilon D_{12}yT_2,$$

so that finally

(27) $$y^*\bar{a}^{n-1}(p^*)\bar{d}_{21}(p^*) = \epsilon d_{12}y.$$

As a consequence of (19),

(28) $$y^*p^* = q^*y^*,$$

so that (27) becomes

(29) $$\bar{a}^{n-1}(q^*)\bar{d}_{21}(q^*)y^* = \epsilon d_{12}y.$$

In the case under consideration, when D_{11} cannot be taken as non-singular, by Lemma V, (29) simplifies to

$$a^{n-1}(p)d_{21}(p)y' = \epsilon d_{12}(p)y$$

or

(30) $$a^{n-1}d_{21}y' = \epsilon d_{12}y.$$

Since $d_{21} = -d_{12}$, this last equation becomes $a^{n-1}y' = -\epsilon y$. When $q = p'$, it is possible to determine y so that $y = y'$ [6] and therefore $a^{n-1} = -\epsilon$.

If the automorphism over K is of period one, $q = p'$ and $a^{n-1} = -\epsilon$, the original matrix D_{11} must also be singular. For, corresponding to (30) we have $a^{n-1}d_{11} = \epsilon d_{11}$ or $d_{11} = -d_{11}$, so that $d_{11} = 0$. Hence we have in place of Lemma V

LEMMA VI. *The matrix D_{11} may be taken as non-singular unless the automorphism over K is of period one, $q = p'$ and $a_1^{n-1}(p) = -\epsilon$. Conversely, if the above conditions are satisfied, the matrix D_{11} is necessarily singular.*

By Lemma VI, if D_{11} may not be taken as non-singular, we may suppose

[6] John Williamson, "The equivalence of non-singular pencils of hermitian matrices in an arbitrary field," *American Journal of Mathematics*, vol. 57 (July, 1933), page 490. Cf. Section 8 of the present paper, formula (44).

that $d_{21} = -d_{12} \neq 0$ and then it is easily shown that $\begin{vmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{vmatrix} \neq 0$. Therefore either $S_{11} \neq 0$ or $\begin{vmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{vmatrix} \neq 0$, so that by Lemma III we may reduce S to one of the forms $[S_{11}, T_{22}]$ or $\left[\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, T_{22}\right]$. We may now repeat the above argument where S is replaced by T_{22}, and finally reduce S to a diagonal block matrix where each block is either of type (i) $\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$ where $|S_{11}| = |S_{22}| = 0$, $\begin{vmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{vmatrix} \neq 0$ or of type (ii) S_{11}, where $|S_{11}| \neq 0$. Hence we need only consider blocks of type (i) or type (ii) separately.

7. Type (i), D_{11} singular. The matrix $\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \sigma_1 + \sigma_2$, where $\sigma_1 = \begin{pmatrix} S_{11} & 0 \\ 0 & S_{22} \end{pmatrix}$ and $\sigma_2 = \begin{pmatrix} 0 & S_{12} \\ S_{21} & 0 \end{pmatrix}$. Since $d_{11} = d_{22} = 0$, σ_1 is nilpotent; since $d_{12} = -d_{21} \neq 0$, σ_2 is non-singular; since

$$[Q_1, Q_2](\sigma_1 + \sigma_2) = (\sigma_1 + \sigma_2)[M_1, M_2], \quad [Q_1, Q_2]\sigma_i = \sigma_i[M_1, M_2], \quad (i = 1, 2).$$

Further

$$[Q_1, Q_2]\sigma_1\sigma_2^{-1} = \sigma_1[M_1, M_2]\sigma_2^{-1} = \sigma_1\sigma_2^{-1}[Q_1, Q_2]$$

and

$$(\sigma_1\sigma_2^{-1})^* = \epsilon^2\sigma_2^{-1}\sigma_1 = \sigma_2^{-1}\sigma_1.$$

Therefore, if

$$R = (E - \tfrac{1}{2}\sigma_1\sigma_2^{-1}), \qquad R[Q_1, Q_2] = [Q_1, Q_2]R,$$

and

$$\begin{aligned} R(\sigma_1 + \sigma_2)R^* &= (E - \tfrac{1}{2}\sigma_1\sigma_2^{-1})(\sigma_1 + \sigma_2)(E - \tfrac{1}{2}\sigma_2^{-1}\sigma_1), \\ &= \sigma_2 - \tfrac{3}{4}\sigma_1\sigma_2^{-1}\sigma_1 + \tfrac{1}{4}\sigma_1\sigma_2^{-1}\sigma_1\sigma_2^{-1}\sigma_1, \\ &= \phi_2 + \phi_1, \end{aligned}$$

where

$$\phi_2 = \sigma_2 - \tfrac{3}{4}\sigma_1\sigma_2^{-1}\sigma_1 \quad \text{and} \quad \phi_1 = \tfrac{1}{4}\sigma_1\sigma_2^{-1}\sigma_1\sigma_2^{-1}\sigma_1.$$

Since $|\phi_2| = |\sigma_2| \neq 0$, ϕ_2 is non-singular and ϕ_1 is nilpotent of index less than that of σ_1. In fact $\phi_1 \equiv 0 \bmod \sigma_1^3$. If the above process be repeated with σ_i replaced by ϕ_i, in a finite number of steps we determine a matrix W such that

$$W[Q_1, Q_2] = [Q_1, Q_2]W \quad \text{and} \quad W\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}W^* = \begin{pmatrix} 0 & L_{12} \\ L_{21} & 0 \end{pmatrix}.$$

Finally, if $X = \begin{pmatrix} E & 0 \\ 0 & L_{21}^{-1} \end{pmatrix}$,

$$X[Q_1, Q_2]X^{-1} = [Q_1, f(Q^*_1)] \quad \text{and} \quad X\begin{pmatrix} 0 & L_{12} \\ L_{21} & 0 \end{pmatrix} X^* = \begin{pmatrix} 0 & \epsilon E_1 \\ E_1 & 0 \end{pmatrix}.$$

Since $[Q_1, f(Q^*_1)]$ is similar to $[Q_1, Q_2]$, by Theorem 1 we may replace the block $[Q_1, Q_2]$ in Q by $[Q_1, f(Q^*_1)]$; the corresponding block in the canonical form of S is $\begin{pmatrix} 0 & \epsilon E_1 \\ E_1 & 0 \end{pmatrix}$. We have therefore

RESULT (b). *Let A be normal with respect to H so that $AH = Hf(A^*)$. If the automorphism over K is of period one, if $f(x) \equiv x \bmod p(x)$ and if $a_1^{n-1}(x) \equiv -\epsilon \bmod p(x)$, where $a_1(x) = f'(x)$, then an elementary factor $[p(x)]^n$ of A must occur an even number of times. Corresponding to each pair of elementary factors $[p(x)]^n$ of A in the canonical form for A and H occur the blocks $[Q, f(Q^*)]$ and $\begin{pmatrix} 0 & \epsilon E \\ E & 0 \end{pmatrix}$, where Q is the Wedderburn canonical form for a matrix with the single elementary factor $[p(x)]^n$.*

The canonical form above is determined uniquely by $f(x)$ and the elementary factor $[p(x)]^n$. Consequently we have

THEOREM 3. *If A and A_1 are two similar matrices, both normal with respect to H, and, if the conditions of result* (b) *are satisfied for every elementary factor of A, then A and A_1 are H-equivalent.*

8. Type (ii), D_{11} non-singular. In type (ii), S_{11} is non-singular and therefore D_{11} is non-singular. Since Q_1 has the single elementary factor $[p(x)]^{e_1}$ the reduction of Q_1 and S_{11} will be the same in principle as that of a matrix Q which has a single elementary factor. To simplify our notation we therefore consider the case in which A and, therefore, Q, has the single elementary factor $[p(x)]^n$. If p is the companion matrix of $p(x)$, Q has the form

$$pE + U.$$

With our previous notation

$$D = D_0 + D_1 + \cdots + D_{n-1},$$

where D_0 is non-singular. Further, since $(pE + U)D = D(pE + \sum_{i=1}^{n-1} a_i(p)U^i)$,

$$UD = D\sum_{i=1}^{n-1} a_i(p)U^i,$$

and in particular

$$UD_0 = a_1 D_0 U. \tag{31}$$

Accordingly,

$$D_0 = [d, da, da, \cdots, da^{n-1}], \quad \text{where} \quad a = a_1(p)$$

and, since $DyT = \epsilon(DyT)^*$,

$$dy = \epsilon(da^{n-1}y)^* = \epsilon y^* d^* (a^{n-1})^*,$$

and

$$ady = \epsilon y^* d^* (a^{n-2})^*.$$

Therefore,

$$ay^*d^*a^* = y^*d^*, \quad \text{or} \quad ay^*a^* = y^*$$

and by (28)

$$a\bar{a}(q^*) = e, \tag{32}$$

where e is the identity matrix of the same order as p.

We now wish to prove

LEMMA VII. *There exists a non-singular matrix R such that*

$$RSR^* = D_0yT \quad \textit{and} \quad R(pE + U)R^{-1} = pE + U + \sum_{i=2}^{n-1} b_i U^i.$$

We shall prove this lemma by induction and accordingly assume that there exists a non-singular matrix W, such that

$$WSW^* = FyT \quad \text{and} \quad W(pE + U)W^{-1} = pE + U + \sum_{i=2}^{r} b_i U^i,$$

where

$$F = F_0 + F_r + F_{r+1} + \cdots + F_{n-1},$$
$$F_0 = D_0, \; F_r \neq 0 \text{ and } F_j \text{ is of the same form as } D_j.$$

Since $QS = SM$,

$$(pE + U + \sum_{i=2}^{r} b_i U^i)F = F(pE + \sum_{i=1}^{n-1} c_i U^i), \tag{33}$$

where the c_i are determined uniquely by the b_i and the polynomial $f(x)$. Consequently

$$b_i U^i F_0 = c_i F_0 U^i, \qquad (i = 1, 2, \cdots, r), \tag{34}$$

and

$$b_1 U F_r = F_r U c_1 + F_0 c_{r+1} U^{r+1}. \tag{35}$$

From (34) and (35) we have

$$b_1 U F_r - c_{r+1}(b_1/c_1)^{r+1} U^{r+1} F_0 = c_1 F_r U$$

and therefore

$$b_1 U G_r + b_{r+1} U^{r+1} G_0 = c_1 G_r U, \tag{36}$$

where

$$G_0 = F_0 = D_0, \quad G_r = -\tfrac{1}{2} F_r, \quad b_{r+1} = \tfrac{1}{2} c_{r+1}(b_1/c_1)^{r+1}.$$

Equation (36) is the same as (16) with D replaced by G and (34) is the same as (15). Therefore by Lemma IV there exists a non-singular matrix

$$G = G_0 + G_r + \cdot\cdot + G_{n-1}, \text{ such that } G \sum_{i=1}^{r} c_i U^i G^{-1} = \sum_{j=1}^{r+1} b_j U^j.$$

Consequently

$$(37) \quad \sum_{j=1}^{r+1} b_j U^j = (GG_0^{-1}) \sum_{i=1}^{r} c_i G_0 U^i G^{-1} = (GG_0^{-1}) \sum_{i=1}^{r} b_i U^i G_0 G^{-1} \text{ by (34)},$$
$$= GG_0^{-1} \sum_{i=1}^{r} b_i U^i (GG_0^{-1})^{-1}.$$

Since $FyT = \epsilon(FyT)^*$, $F_i yT = \epsilon y^* TF^*_i$ and therefore

$$yTF_0^{-1*}F^*_r = F_0^{-1}F_r yT.$$

But

$$GG_0^{-1} = E - \tfrac{1}{2}F_r F_0^{-1} + G_{r+1}F_0^{-1} + \cdot\cdot\cdot + G_{n-1}F_0^{-1}$$

and consequently

$$yT(GG_0^{-1})^* = (E - \tfrac{1}{2}F_0^{-1}F_r + L_{r+1} + \cdot\cdot\cdot + L_{n-1})yT,$$

where L_i is of the same form as D_i. Hence

$$(38) \qquad GG_0^{-1}FyT(GG_0^{-1})^* = (F_0 + X_{r+1} + \cdot\cdot\cdot + X_{n-1})yT,$$

where X_i is of the same form as D_i. If $J = GG_0^{-1}W$, it follows from (36) and (38) that

$$JSJ^* = (D_0 + X_{r+1} + \cdot\cdot\cdot + X_{n-1})yT$$

and

$$J(pE + U)J^{-1} = pE + U + \sum_{i=2}^{r+1} b_i U^i.$$

Hence, by induction on r the lemma is proved. Accordingly in the canonical form for A and H we may take

$$(39) \quad Q = pE + U + \sum_{i=2}^{n-1} b_i U^i, \quad S = D_0 yT, \quad M = qE + \sum_{j=1}^{n-1} c_j(q) U'^j,$$

where $D_0 = [d, ad, a^2d, \cdot\cdot\cdot, a^{n-1}d]$ and $a = f'(p)$. The matrices b_j and d are polynomials in p and the c_j are determined completely by the b_i and $f(x)$. The matrices b_j and c_j satisfy certain simple equations, which we proceed to determine.

In place of (34) we have

$$(40) \qquad b_i U^i D_0 = c_i D_0 U^i, \qquad (i = 1, 2, \cdot\cdot\cdot, n-1),$$

from which we get when $i = 1$,

$$UD_0 = c_1 D_0 U = aD_0 U,$$

and therefore

(41) $$c_i = b_i a^i, \qquad (i = 1, 2, \cdots, n-1),$$

so that in particular $c_1 = a$.

From equation (36) we obtain by multiplying throughout by yT

(42) $$b_{r+1} U^{r+1} G_0 yT = (aG_r U - UG_r) yT.$$

Now

$$\begin{aligned}\{(aG_rU - UG_r)yT\}^* &= y^*T(\bar{a}(p^*)U'G^*_r - G^*_rU'),\\ &= \epsilon\{\bar{a}(q^*)UG_r - G_rU\}yT,\\ &= \epsilon a^{-1}\{UG_r - aG_rU\}yT \text{ by } (32).\end{aligned}$$

Therefore, by (42),

$$\begin{aligned}-\epsilon a^{-1} b_{r+1} U^{r+1} G_0 yT &= (b_{r+1}U^{r+1}G_0yT)^*\\ &= y^*T\bar{b}_{r+1}(p^*)G^*_0U'^{r+1},\\ &= \epsilon\bar{b}_{r+1}(q^*)G_0U^{r+1}yT,\\ &= \epsilon\bar{b}_{r+1}(q^*)a^{-r-1}U^{r+1}G_0yT.\end{aligned}$$

Consequently

(43) $$\begin{cases}\bar{b}_{r+1}(q^*) = -a^r(p)b_{r+1}(p) = -a^{-1}(p)c_{r+1}(p)\\ \bar{b}_{r+1}(p^*) = -a^r(q)b_{r+1}(q) = -a^{-1}(q)c_{r+1}(p).\end{cases}$$

Since D_0yT is hermitian or anti-hermitian the matrix d satisfies

$$dy = \epsilon y^*\bar{d}(p^*)\bar{a}(p^*)^{n-1}.$$

Further, since $py = yq$,

$$y^*p^* = q^*y^* \text{ and } y^*\bar{f}(p^*) = \bar{f}(q^*)y^* \text{ or by (3) } y^*q = py^*.$$

Accordingly

(44) $$y^* = h(p)y,$$

and

$$y = y^*\bar{h}(p^*) = h(p)y\bar{h}(p^*) = h(p)\bar{h}(q^*)y, \text{ so that}$$

(45) $$h(p)\,\bar{h}(q^*) = e$$

and finally

(46) $$d = \epsilon h(p)\bar{d}(q^*)\,\bar{a}(q^*)^{n-1}.$$

We have therefore proved

Result (c). *Let A be normal with respect to H, so that $AH = Hf(A^*)$, and let A have the single elementary factor $[p(x)]^n$. Then, if p is the companion matrix of $p(x)$, a canonical form for A and H is given by* (39). *The matrices a, b_i, c_j, d and y satisfy equations* (32), (43), (44), (45) *and* (46).

For the first time the canonical form is not completely determined by $f(x)$ and the elementary factor $[p(x)]^n$, since the matrix $d(p)$ is independent

both of $f(x)$ and $p(x)$. If the elementary factors of A are $[p(x)]^n$ repeated exactly s times, a canonical form for H would involve exactly s matrices $d_1(p)$, $d_2(p), \cdots, d_s(p)$, one for each factor $[p(x)]^n$. We have therefore proved,

THEOREM 4. *If A is normal with respect to H, so that $AH = Hf(A^*)$, a canonical form for A, H is Q, S, where Q and S are direct sums of matrices given by results* (a), (b) *and* (c). *In general the canonical form is not uniquely determined by $f(x)$ and the invariant factors of A.*

COROLLARY. *Two similar matrices both normal with respect to the same hermitian or anti-hermitian matrix are not necessarily H-equivalent.*

9. Necessary and sufficient conditions for *H*-equivalence. If A is normal with respect to H, there may exist two distinct canonical forms Q, S_1 and Q, S_2 for A, H. If this is so, by theorem 1 there must exist a non-singular matrix R such that

$$RQ = QR \text{ and } RS_1R^* = S_2.$$

If Q is a diagonal block matrix, where each block has a single distinct characteristic factor, since R is commutative with Q, R is a diagonal block matrix partitioned similarly to Q. Accordingly it is sufficient to suppose that Q has a single characteristic factor and has the form used in § 6. Further, as a consequence of result (b), we need only consider the case in which

$$S_1 = DyT = [D_1yT_1,\ D_2yT_2, \cdots, D_kyT_k] \quad \text{and}$$
$$S_2 = FyT = [F_1yT_1,\ F_2yT_2, \cdots, F_kyT_k],$$

where

$$D_i = [d_i, ad_i, \cdots, a^{e_i-1}d_i] \text{ and } F_i = [f_i, af_i, \cdots, a^{e_i-1}f_i].$$

Since

$$RS_1R^* = S_2,$$

$$RD\bar{R} = F, \tag{47}$$

where

$$\bar{R} = yTR^*(yT)^{-1}.$$

Since $RQ = QR$, $R^*Q^* = Q^*R^*$ and $yTQ^*(yT)^{-1}\bar{R} = \bar{R}yTQ^*(yT)^{-1}$. But $yTQ^*(yT)^{-1}$ is obtained from Q by replacing p by q^*, so that $\bar{R}$ is of the same general form as R. If $R = (R_{ij})$ and $\bar{R} = (\bar{R}_{ij})$, $i, j = 1, 2, \cdots, k$, (47) becomes

$$\sum_{a=1}^{k} R_{ia}D_a\bar{R}_{aj} = \delta_{ij}F_i, \qquad (i, j = 1, 2, \cdots, k), \tag{48}$$

where δ_{ij} is the Kronecker delta.

Let r_{ij} and $\bar{r}_{ij}$ denote the elements in the first row and first column of R_{ij} and $\bar{R}_{ij}$ respectively. Then from the nature of the matrices R_{ij}, $\bar{R}_{ij}$ it follows from (48) that

$$\sum_{a=1}^{k} r_{ia}d_a\bar{r}_{aj} = \delta_{ij}f_i, \qquad (i, j = 1, 2, \cdots, k). \tag{49}$$

If $e_{c-1} > e_c = e_{c+1} = \cdots e_d > e_{d+1}$,

$r_{ia} = 0$, if $\alpha < c$, $\bar{r}_{aj} = 0$, if $\alpha > d$, for all $i, j = c, c+1, \cdots, d$. Hence (49) yields

$$\sum_{a=c}^{d} r_{ia}d_a\bar{r}_{aj} = f_i\delta_{ij}, \qquad (i, j = c, c+1, \cdots, d),$$

or

$$\sum_{a=c}^{d} r_{ia}(p)d_a(p)\bar{r}_{ja}(q^*) = \delta_{ij}f_i(p), \qquad i, j = c, \cdots, d). \tag{50}$$

By re-arrangement of the rows and columns of R, it may be shown that $| r_{ij} |$, $i, j = c, c+1, \cdots, d$, is a factor of $| R |$ and is therefore non-zero.

Let θ be a zero of $p(x)$ and Ω be the simple algebraic adjunction field $K(\theta)$. Since $q^* = \bar{f}(p)$ and satisfies $\bar{p}(x) = 0$, if $\phi = \bar{f}(\theta)$ and $g(\theta)$ is any element of Ω, the correspondence $g(\theta) \leftrightarrow \bar{g}(\phi)$ induces an automorphism of Ω. This automorphism is of period one or two and we can therefore define the conjugate transposed of a matrix with elements in Ω. Equation (49) is equivalent to

$$\sum_{a=c}^{d} r_{ia}(\theta)d_a(\theta)\bar{r}_{ja}(\phi) = \delta_{ij}f_i(\theta), \qquad (i, j = c, c+1, \cdots, d),$$

and this last may be written more compactly in the form

$$\mathrm{P}[d_c(\theta), d_{c+1}(\theta), \cdots, d_d(\theta)]\mathrm{P}^* = [f_c(\theta), \cdots, f_d(\theta)], \tag{51}$$

where $\mathrm{P} = (r_{ij}(\theta))$, $i, j = c,\ c+1, \cdots, d$, and P^* is the conjugate transposed of P. Conversely, if (51) is satisfied, so is (50) and the equations obtained from (50) by multiplying by a^s, $s = 1, \cdots, e_c - 1$. Consequently,

$$\sum_{a=c}^{d} r_{ia}D_a\bar{r}_{ja}(q^*) = \delta_{ij}F_i, \qquad (i, j = c, c+1, \cdots, d),$$

and finally

$$W[D_cyT_c, D_{c+1}yT_{c+1}, \cdots, D_dyT_d]W^* = [F_cyT_c, F_{c+1}yT_{c+1}, \cdots, F_dyT_d],$$

where $W = (r_{ij}E_c)$, $i, j = c,\ c+1, \cdots, d$. Since $| \mathrm{P} | \neq 0$, W is non-singular.

If there are t distinct powers of $p(x)$ that occur among the elementary factors of A, there will be exactly t equations (51). These t equations are the necessary and sufficient conditions that the two canonical forms be equivalent. Let us for convenience call the diagonal matrix $[d_c(\theta), \cdots, d_d(\theta)]$ *the diagonal matrix associated* with the elementary factor $[p(x)]^{e_c}$. We therefore have

THEOREM 5. *When A is normal with respect to H, two canonical forms for A and H are equivalent if, and only if, the diagonal matrices associated with each elementary factor $[p_i(x)]^{e_{ij}}$ of A are conjunctively equivalent in the field $K(\theta_i)$, where θ_i is a zero of $p_i(x)$.*

THEOREM 6. *When A and A_1 are similar matrices both normal with respect to H, they are H-equivalent if, and only if, for all values of i and j, the diagonal matrix associated with the elementary factor $[p_i(x)]^{e_{ij}}$ in a canonical form for A and H is conjunctively equivalent in $K(\theta_i)$ to that of the canonical form for A_1 and H.*

THEOREM 7. *Let A_1 be normal with respect to H_1 and A_2 normal with respect to H_2. Necessary and sufficient conditions that there exist a non-singular matrix P such that $PA_1P^{-1} = A_2$ and $PH_1P^* = H_2$, are (i) the invariant factors of A_1 be the same as those of A_2 and (ii) for all values of i and j the diagonal matrix associated with $[p_i(x)]^{e_{ij}}$ in a canonical form for A_1 and H_1 be conjunctively equivalent in $K(\theta_i)$ to that of a canonical form for A_2 and H_2.*

10. Particular fields *K*. Certain simplifications arise when K is the field of all complex numbers and A^* is the conjugate transposed of A in the classical sense. Results (a) and (b) are practically unaltered but result (c) can be greatly simplified. The field K is algebraically closed, so that p is a matrix of one row and column, $q^* = p$ and $\Omega = K(\theta) = K$. Equation (32) reduces to $a\bar{a} = 1$ so that $a = e^{i\omega}$ and (46) yields $d = \epsilon\bar{d}e^{(n-1)i\omega}$. If $d = \tau e^{i\phi}$, $e^{2i\phi} = \epsilon e^{(n-1)i\omega}$ and for a fixed value of $\sqrt{\epsilon e^{(n-1)i\omega}} = \alpha$, $e^{i\phi} = \pm\alpha$. Therefore $d = \pm\tau\alpha$ where α is uniquely determined. The matrix $\mathbf{P}^*$ in (51) is the conjugate transposed of $\mathbf{P}$ and, if $d_i = \delta_i\alpha$, and $f_i = \psi_i\alpha$, where δ_i and ψ_i are real, (51) states that $[\delta_c, \delta_{c+1}, \cdots, \delta_d]$ have the same index as $[\psi_c, \psi_{c+1}, \cdots, \psi_d]$. This final result gives Theorem 4 of the paper dealing with this special case. The somewhat simpler canonical form found previously for this particular case can now be obtained by a simple transformation.[7]

If K is the field of all real numbers there are two distinct types of irreducible factors $p(x)$, either linear or quadratic. If $p(x)$ is linear, (51) shows that $[d_c, \cdots, d_d]$ must have the same index as $[f_c, \cdots, f_d]$. If $p(x)$ is quadratic, $K(\theta)$ is the field of all complex numbers. If $\phi = \theta$, so that the automorphism over Ω is of period one, the matrix $\mathbf{P}^*$ is the transposed of $\mathbf{P}$. Since $\Omega = K(\theta)$ is the field of all complex numbers, there always exists a non-singular matrix $\mathbf{P}$ satisfying (51) and the conditions imposed by (51) may be omitted. In the remaining case when $\phi = \bar{\theta}$, the conjugate of θ, the automorphism over Ω is of period two and $\mathbf{P}^*$ is the conjugate transposed of $\mathbf{P}$.

[7] See footnote (1).

The matrix a is an orthogonal matrix of order 2, y is the identity matrix and $q^*_a = p'$. If instead of writing $p = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ the notation $a + ib$ is adopted, the arguments used above, for K the field of all complex numbers, may be repeated word for word and again (51) is equivalent to the fact that the indices of two real diagonal matrices be the same.

If K is the field of all complex numbers and the automorphism over K is of period one, the fact that $K(\theta) = K$ and $\mathbf{P}^* = \mathbf{P}'$ shows that the conditions imposed by (51) may be omitted.

The exact conditions imposed by (51) for a general field K are not known, but they depend on the arithmetic properties of the field.[8]

11. Particular cases of $f(x)$. Some particularizations of the polynomial $f(x)$ are of special interest. The first of these is when $f(x) = x$. Then $AH = HA^*$, so that

$$(AH)^* = H^*A^* = \epsilon HA^* = \epsilon AH,$$

and the matrix AH is also hermitian or anti-hermitian. Further, if

$$PAP^{-1} = A_1 \text{ and } PAP^* = H,$$
$$PAHP^* = A_1H_1.$$

Therefore by particularizing $f(x)$ to be x, the results of this paper may be applied to the conjunctive equivalence of two pencils of non-singular hermitian or anti-hermitian matrices. If $\epsilon = -1$, and the automorphism over K is of period one, both matrices AH and H are skew symmetric. Since $f'(x) = +1$, $a = +1$ and for all values of n, $a^{n-1} = -\epsilon$. Therefore the conditions of result (b) are always satisfied, and, as an immediate consequence of theorems 2 and 3, we have

THEOREM 8. *Two non-singular pencils of skew symmetric matrices in an arbitrary field are congruently equivalent, if, and only if, the two pencils have the same invariant factors.*[9]

Let $f(x) = x$ and $\epsilon = 1$. If $p_i(x)$ is a characteristic factor of A and $\bar{p}_i(x) \neq p_i(x)$, the matrix p^*_i satisfies $\bar{p}_i(x)$ and $p_i(x)$ is a characteristic factor of type (i). A characteristic factor of type (ii) is a polynomial $p(x)$

[8] L. E. Dickson, "On quadratic, bilinear and hermitian forms," *Transactions of the American Mathematical Society*, vol. 7 (1906), pp. 275-292; "On quadratic forms in a general field," *Bulletin of the American Mathematical Society*, vol. 14 (1907-8), pp. 108-115. H. Hasse, "Symmetrische Matrizen in Körper der rationalen zahlen," *Crelle*, vol. 153, pp. 12-43.

[9] Cf. John Williamson, "The conjunctive equivalence of pencils of hermitian and anti-hermitian matrices," *American Journal of Mathematics*, vol 59 (April, 1937), No. 2, pp. 399-413, Theorem 7.

which is unaltered by the automorphism over K. Since $f'(x)=1$, $a=1$, $q^*=p$ and, as stated earlier, $y=y'$ and (46) yields $d=\bar{d}$. Equations (43) now become $\bar{b}_{r+1}=-b_{r+1}=-c_{r+1}$. Since $Q^*=f(Q^*)$, $\bar{b}_{r+1}=c_{r+1}$ and accordingly $b_{r+1}=0$, $r=1,2,\cdots,n-2$. Equation (51) is not simplified beyond the fact that the elements $d_i(\theta)$, $f_i(\theta)$ are unaltered by the original automorphism over K. From result (a) and the simplified form of (39) we obtain a canonical form Q, S for the matrices A, H and therefore a canonical form QS, S for the pair of hermitian matrices AH, H under non-singular conjunctive transformations. The canonical forms thus obtained differ slightly from those obtained in a previous paper,[10] as previously the matrices were considered as matrices not over K but over the field unaltered by the automorphism over K.

In a similar manner, if $f(x)=-x$ and $\epsilon=-1$, the conditions for the equivalence of two pairs of matrices, one symmetric and the other skew symmetric and non-singular, under congruent transformations may be deduced.[11]

If $f(A^*)=(A^*)^{-1}$, AHA^* is a conjunctive automorph of H and the problem under consideration reduces to that of the equivalence of two conjunctive automorphs of H under similarity transformations by matrices, which are also conjunctive automorphs of H. This special problem is the simple extension of that of the unitary equivalence of two unitary matrices. The particular case of this problem when K is the complex field and $\epsilon=1$ [12] has been dealt with, as well as that in which K is the real field and $\epsilon=-1$.[13]

12. H singular. It seems unlikely that the above results could be extended to the case in which H is singular without some further hypotheses. If, for example, H is the zero matrix, the fact that A is normal with respect to H is vacuous, inasmuch as every matrix is now normal with respect to H. If A is normal with respect to H and H is non-singular, the correspondence between A and $f(A^*)$ is a particular case of an involution [14] and by (3) $f\{\bar{f}(A)\}=A$. When H is singular a practical definition of normality would be: *A matrix A is normal with respect to H if $AH=Hf(A^*)$ and $f\{\bar{f}(A)\}=A$.* With this definition of normality, it would be possible to extend the results of this paper to the case in which H is singular. Certain complications arise,

[10] See 6 page 487.

[11] See John Williamson, "On the algebraic problem concerning the normal forms of linear dynamical systems," *American Journal of Mathematics*, vol. 58 (January, 1936), pp. 141-163.

[12] John Williamson, "Quasi-unitary matrices," *Duke Mathematical Journal*, vol. 3 (December, 1937), No. 4, pp. 715-25.

[13] John Williamson, "On the normal forms of linear canonical transformations," *American Journal of Mathematics*, vol. 59 (July, 1937), No. 3, pp. 599-617.

[14] Cf. A. A. Albert, *Modern Higher Algebra*, University of Chicago Press, 1937, Chapter V.

for example in the proof of result (a); S_{ij} is non-singular and the form of S_{ij} would have to be considered in more detail. If this is done, it can be shown that result (a) still holds, if E is replaced by U^r, where $n-r$ is the rank of S_{ij}. For the sake of brevity, the remaining results will not be given here.

13. Characteristic of *K* different from zero. If the characteristic of K is different from two, all results hold except when $p(x)$ is inseparable. Then the usual Wedderburn canonical form for a matrix with the elementary factor $[p(x)]^{e_i}$ is no longer available. However, we may use instead the matrix [15]

$$Q_i = pE + mU, \tag{52}$$

where every element of m is zero except the element in the last row and column, which has the value unity. The coefficient of U in $Q_i{}^{t+1}$ is

$$\sum_{j=0}^{t} p^j m p^{t-j} \text{ and } p\sum_{j=0}^{t} p^j m p^{t-j} - \sum_{j=0}^{t} p^j m p^{t-j} p = p^{t+1}m - mp^{t+1}.$$

Therefore, if h is the coefficient of U in $p(Q_i)$,

$$ph - hp = p(p)m - mp(p) = 0,$$

and h is commutative with p. Hence h is a polynomial in p so that we may replace m in (52) by $h^{-1}p$ or, what is equivalent to this, assume that the coefficient of U in $p(Q_i)$ is unity. With the notation of § 6

$$M_i = qE + \sum r_j U'^j,$$

where the coefficient a of U' in $p(M_i)$ is a polynomial in q. If $S = DyT$, each D_{ij} satisfies the conditions of lemma III except that the elements of D_{ij} are not all polynomials in p. Since $p(Q_i)S = Sp(M_i)$, lemma VI is true where

$$D_0 = [d, ad, \cdots, a^{n-1}d]$$

and d is a polynomial in p. Unfortunately, at present we have not been able to determine the form that Q assumes when D is reduced to diagonal form. The use of the alternative canonical form given by Wedderburn [15] did not lead to any more satisfactory results. Accordingly the problem when A has a single elementary factor $[p(x)]^n$ and $p(x)$ is inseparable is as yet unsolved.

The case of characteristic two has not been considered. This is always an exceptional case and certainly the methods of this paper, which use so often division by two, would not apply. It is hoped that the methods used by Albert [16] may be applicable and lead to a satisfactory solution.

THE JOHNS HOPKINS UNIVERSITY.

---

[15] J. H. M. Wedderburn, "The canonical form of a matrix," *Annals of Mathematics*, vol. 39 (January, 1938), No. 1, pp. 178-80.

[16] A. A. Albert, "Symmetric and alternate matrices in an arbitrary field," *Transactions of the American Mathematical Society*, vol. 43 (May, 1938), No. 3, pp. 386-436.

THE WARING PROBLEM WITH SUMMANDS $1 + bx^n$.*

By Mary Haberzetle.

A positive integer will be said to be represented by $[u, w]$ if it is a sum of u positive integers of the form $1 + bx^n$ or 0 and w positive integers of the form $a(1 + bx^n)$ or 0, where a and b are fixed positive integers. In this paper u and w will be determined such that for $19 \leqq n \leqq 400$, $1 \leqq a \leqq 4n$, $1 \leqq b \leqq 2n + 1$ every positive integer is represented by $[u, w]$. The problem is a modified Waring problem and was suggested by L. E. Dickson. The results obtained are stated in Theorem 3.

For brevity define $p_0 = 0$, $p_1 = 1$, $p_2 = 1 + b$, $p_3 = 1 + b \cdot 2^n, \cdots, p_x = 1 + b(x-1)^n$. For $v > 0$ denote by v' the number of terms, 0, 1, $1 + b$, required to represent v, and by v'' the number of terms, 0, 1, $1 + b$, required to represent each of $1, \cdots, v$. For $v = 0$, define $v' = v'' = 0$. The notation $[x]$ is used to denote the integral part of x.

1. Lemmas for ascent. Lemmas $1, \cdots, 4$ have been proved by L. E. Dickson.[1]

Lemma 1. *If $z \geqq a$ and if all integers in the interval $(E, E + zp_k)$ are represented by $[A, B]$, then all in $(E, E + (z + av)p_k)$ are represented by $[A, B + v]$.*

Lemma 2. *If $z \geqq 1$ and all integers in $(E, E + zp_k)$ are represented by $[A, B]$, then all in $(E, E + (z + w)p_k)$ are represented by $[A + w, B]$.*

Lemma 3. *Let R_j denote the least positive or zero residue modulo a of $G_j = [p_j/p_{j-1}]$, $j > 1$. Let $Q_j = (G_j - R_j)/a$. If all integers in $(E, E + ap_{m-1})$ are represented by $[R, S]$, then all in $(E, E + p_m)$ are represented by $[R, S + Q_m]$.*

Lemma 4. *If $m \geqq 4$ and if all integers in $(E, E + ap_3)$ are represented by $[A, B]$, then all in $(E, E + p_m)$ are represented by $[C, D]$, $C = A + (m-4)(a-1)$, $D = B + Q_4 + \cdots + Q_m$.*

Lemma 5. *Let C be a positive integer and define c to be the least positive residue of C modulo a. Each of the integers $0, 1, \cdots, C - 1$ is represented by*

* Received June 9, 1938.

[1] "Universal forms $\Sigma a_i x_i^n$ and Waring's problem," *Acta Arithmetica*, vol. 2 (1937), pp. 178-179.

$$[(c-1)'', k'] \quad or \quad [(a-1)'', (k-1)''],$$

where $k = (C-c)/a$. *The second form is to be suppressed if* $k = 0$.

Those integers which are $\geqq C - c = ka$ are sums of ka and $0, 1, \cdots, c-1$. The latter are sums of $(c-1)''$ terms $0, 1, 1+b$, and ka is a sum of k' terms $0, a, a(1+b)$.

Only when $k \geqq 1$ are there further integers, $ka - ma + j$ $(j = 0, 1, \cdots, a-1;\ k \geqq m \geqq 1)$, and these are represented by $[(a-1)'', (k-1)'']$.

Lemma 6. *Let* $p_4 = qp_3 + r$, $0 \leqq r < p_3$. *Denote by* f, g, *and* d *the least positive residues modulo* a *of* $p_3 - r$, r, *and* q, *respectively. Then every integer in the interval* $(qp_3, (q+1)p_3)$ *is represented by one of*

$$\text{(1)} \qquad \left[d+1+(g-1)'', \left(\frac{r-g}{a}\right)' + \left(\frac{q-d}{a}\right)\right],$$

$$\text{(2)} \qquad \left[d+(a-1)'', \left(\frac{r-g}{a}-1\right)'' + \left(\frac{q-d}{a}\right)\right],$$

$$\text{(3)} \qquad \left[(f-1)''+1, \left(\frac{p_3-r-f}{a}\right)'\right],$$

$$\text{(4)} \qquad \left[(a-1)''+1, \left(\frac{p_3-r-f}{a}-1\right)''\right].$$

The second is to be suppressed if $r \leqq a$ *and the fourth if* $p_3 - r \leqq a$.

By Lemma 5, $0, 1, \cdots, r-1$ are represented by $\left[(g-1)'', \left(\frac{r-g}{a}\right)'\right]$ if $r \leqq a$, but by it or $\left[(a-1)'', \left(\frac{r-g}{a}-1\right)''\right]$ if $r > a$. To these add $\left[d, \frac{q-d}{a}\right]$ which represents $qp_3 = dp_3 + \frac{q-d}{a} \cdot ap_3$. The weight of (1) is increased by one for reasons which will appear shortly.

The remaining integers in the interval (except the final one) are sums of $qp_3 + r = p_4$ and $0, 1, \cdots, p_3 - r - 1$. These are represented by $\left[(f-1)''+1, \left(\frac{p_3-r-f}{a}\right)'\right]$ or $\left[(a-1)''+1, \left(\frac{p_3-r-f}{a}-1\right)''\right]$.

Finally, $(q+1)p_3$ is represented by $\left[d+1, \frac{q-d}{a}\right]$ and hence by (1).

Lemma 7. *Every integer in the interval* $(qp_3, (q+a)p_3)$ *is represented by one of the forms obtained by adding* $a-1$ *to the first entries of the forms* (1), $\cdots$, (4).

The lemma follows from Lemmas 2 and 6.

2. Representation of integers $\leqq qp_3-1$.

LEMMA 8. *Let $a<p_3$. Let c and d be the least positive residues modulo a of p_3 and q, respectively. Then every positive integer $\leqq qp_3-1$ is represented by one of the two forms,*

$$[U_0,V_0]=\left[a-1+(c-1)'',\left(\frac{p_3-c}{a}\right)'+\frac{q-d}{a}\right],$$

$$[U_1,V_1]=\left[a-1+(a-1)'',\left(\frac{p_3-c}{a}-1\right)''+\frac{q-d}{a}\right].$$

The integers in question are xp_3+y, $0\leqq x\leqq q-1$, $0\leqq y\leqq p_3-1$. By Lemma 5 the y's are represented by

$$\left[(c-1)'',\left(\frac{p_3-c}{a}\right)'\right]\quad\text{or}\quad\left[(a-1)'',\left(\frac{p_3-c}{a}-1\right)''\right].$$

Since $q-1\equiv d-1 \pmod a$ and $d\leqq a$, xp_3 is represented by $[a-1, (q-d)/a]$. Hence all are represented by $[U_0,V_0]$, $[U_1,V_1]$.

THEOREM 1. *Let $1\leqq a\leqq 4n$, $1\leqq b\leqq 2n+1$. Let c and d be the least positive residues modulo a of p_3 and q, respectively. Then every positive integer $\leqq qp_3-1$ is represented by*

$$H=\left[4n+a(2n+1)+(a-1)''-1,\right.$$
$$\left.\left(\frac{p_3-c}{a}\right)''+\frac{q-d}{a}-[4n/a]-2n\right]. \tag{5}$$

In Lemma 8 increase U_0 to U_1. Both $\left(\frac{p_3-c}{a}\right)'$ and $\left(\frac{p_3-c}{a}-1\right)''$ are less than or equal to $\left(\frac{p_3-c}{a}\right)''$. It is a consequence then of Lemma 8 that every positive integer $\leqq qp_3-1$ is represented by

$$F=\left[a-1+(a-1)'',\left(\frac{p_3-c}{a}\right)''+\frac{q-d}{a}\right].$$

In F replace each of $[4n/a]+2n$ terms ap_x by $p_x+\cdots+p_x$ (with a summands) to obtain H. It will be shown later that

$$\left(\frac{p_3-c}{a}\right)''+\frac{q-d}{a}-[4n/a]-2n\geqq 0. \tag{6}$$

3. Formulas for t ascents at one step. No proofs will be given for the lemmas in this section. References are made to similar lemmas and their proofs.

LEMMA [2] 9. *If* $g \geqq 0$ *and* $s \geqq g + a$, *there exists a positive integer* i *such that*

$$g \leqq s - a(1 + bi^n) < g + n(abs^{n-1})^{1/n}.$$

LEMMA [3] 10. *Let* g *be an integer* $\geqq 0$, L *an integer* $\geqq a$,

$$G = \{(ab)^{-1}(L/n)^n\}^{1/(n-1)}.$$

If $G \geqq g + L$ *and if all integers between* g *and* $g + L$ *inclusive are represented by* $F = [R, S]$, *then every integer* s *between* g *and* G *inclusive is represented by* $F_1 = [R, S + 1]$.

Write $L_0 = g + L$, $v = (1 - g/L_0)/n$, $L_1 = G$. Lemma 10 becomes

LEMMA 11. *Let* g *be an integer* $\geqq 0$, $L_0 \geqq g + a$,

$$L_1 = \{(ab)^{-1}(L_0 v)^n\}^{1/(n-1)}.$$

If $L_1 \geqq L_0$ *and if all integers between* g *and* L_0 *inclusive are represented by* $F = [R, S]$, *then all between* g *and* L_1 *inclusive are represented by* $F_1 = [R, S + 1]$.

LEMMA [4] 12. *Let* g *be an integer* $\geqq 0$, $v = (1 - g/L_0)/n$, $L_0 \geqq g + a$, $L_0 v^n \geqq ab$. *Compute* L_t *from*

$$\log_{10} L_t = \left(\frac{n}{n-1}\right)^t (\log_{10} L_0 + w) - w,$$

$w = n \log_{10} v - \log_{10} ab$. *If all integers between* g *and* L_0 *inclusive are represented by* $F = [R, S]$ *then all between* g *and* L_t *inclusive are represented by* $F_t = [R, S + t]$.

Apply the above lemma with $a = 1$ and L_t in place of L_0 to obtain

LEMMA 13. *Let* g *be an integer* $\geqq 0$, $V = (1 - g/L_t)/n$, $L_t \geqq g + 1$, $L_t V^n \geqq b$. *If all integers between* g *and* L_t *inclusive are represented by* $F_t = [R, S + t]$ *then all between* g *and* L_{t+T} *are represented by* $F_{t+T} = [R + T, S + t]$ *and*

$$\log_{10} L_{t+T} = \left(\frac{n}{n-1}\right)^{t+T} (\log_{10} L_0 + w) + \left(\frac{n}{n-1}\right)^T (W - w) - W, \tag{7}$$

where $W = n \log_{10} V - \log_{10} b$.

---

[2] L. E. Dickson, "Generalizations of Waring's theorem on fourth, sixth, and eighth powers," *American Journal of Mathematics*, vol. 49 (1927), p. 242.

[3] Dickson, "Universal forms $\Sigma a_i x_i^n$ and Waring's problem," *op. cit.*, pp. 184-185.

[4] *Ibid.*, p. 185.

Lemma 13 will be used for L_t so large that $1/n$ is a sufficiently close approximation to V, whence $W = -n \log_{10} n - \log_{10} b$. But $v < 1/n$, whence $W - w > \log_{10} a$. Since the last two terms of (7) are positive

$$(8) \qquad \log_{10} L_{t+T} > \left(\frac{n}{n-1}\right)^{t+T} (\log_{10} L_0 + n \log_{10} v - \log_{10} ab).$$

4. Results from the asymptotic theory and prime number theory. The following theorem is a consequence of a lemma [5] proved by L. E. Dickson.

THEOREM 2. *Let k_0 be the least integer exceeding $(\log r_1)/(-\log(1-v))$, where $r_1 = n^6(6n-1)/(n-d)$, $d = 1 + 2n^2 z$, $z = v^3/12$ and $v = 1/n$. Take $k_1 = 2k_0$. Then if $1 \leqq a \leqq 4n$, $n \geqq 9$, every integer $\geqq N_1$ is represented by $[4n + b - 1, 3k_1 - 2]$ where $N_1 = bN + 4n + a(3k_1 - 2)$ and $\log_{10} N = .8n^6$.*

LEMMA [6] 14. *If i is a positive integer and $K_i = [i^n/(i-1)^n]$, $m \geqq 6$, $n \geqq 9$, then $\sum_{i=4}^{m} K_i < X$, $X = K_4 + 5K_5 + m - n - 9 + n \log(m-1)/4 - \binom{n}{2}/(m-1)$.*

5. Solution for $19 \leqq n \leqq 400$.

THEOREM 3. *Let $1 \leqq a \leqq 4n$ and $1 \leqq b \leqq 2n + 1$. Let $19 \leqq n \leqq 400$ and denote by c and d the least positive residues modulo a of p_3 and q, respectively. Then every positive integer is represented by*

$$H = \left[4n + a(2n+1) + (a-1)'' - 1, \quad \left(\frac{p_3 - c}{a}\right)'' + \frac{q-d}{a} - [4n/a] - 2n\right].$$

For integers $\leqq qp_3 - 1$ the theorem is true by Theorem 1. For integers $\geqq N_1$ it is true by Theorem 2 since

$$b \leqq 2n + 1 \leqq a(2n+1) + (a-1)'',$$

and

$$(9) \qquad \left(\frac{p_3 - c}{a}\right)'' + \frac{q-d}{a} - [4n/a] - 2n \geqq 3k_1 - 2, \qquad n \geqq 19.$$

The latter inequality is obtained by noting that

$$\left(\frac{p_3 - c}{a}\right)'' \geqq (2^n + 1)/2a - 1/2 \quad \text{and} \quad k_1 \leqq 4n \log n + 2n \log 6.$$

[5] *Ibid.*, p. 190.

[6] *Ibid.*, p. 188.

As a consequence of these inequalities (9) is implied by

$$1 \geqq (96n^2 \log n + 48n^2 \log 6 + 8n + 16n^2)/2^n. \tag{10}$$

For $n = 17$, (10) is true, and since the right-hand side decreases as n increases it holds for $n \geqq 17$.

Since $3k_1 - 2 > 0$, it is now seen that (6) is true.

It remains then to prove Theorem 3 for integers $\geqq qp_3$ and $< N_1$. This is done by ascent. From Lemma 4 and the first part of Lemma 7 all integers in the interval $J = (qp_3, qp_3 + q_{m+1})$ are represented by $[K, D]$, where $K = A + (m-2)(a-1)$ and $D = B + \sum_{i=4}^{m+1} Q_i$, while $[A, B]$ was initially one of the forms (1), $\cdots$, (4) but may be replaced by one of

$$\left[d + 1 + (a-1)'', \left(\frac{r-g}{a}\right)'' + \frac{q-d}{a}\right], \quad \left[(a-1)'' + 1, \left(\frac{p_3 - r - f}{a}\right)''\right]. \tag{11}$$

Apply Lemmas 12, 13 with $g = p_4$, $L_0 = p_{m+1}$, both in J. The conditions, $L_0 v^n \geqq ab$, $L_t V^n \geqq b$, will later be seen to hold after m is restricted. Let Z be the least integral value of $t + T$ for which the second member of (8) exceeds $\log_{10} N_1$, where N_1 is the constant in Theorem 2. It will be shown that $[T + K, t + D] \leqq H$.

For the first case from (11), let

$$T = 4n + a(2n+1) - d - 2 - (a-1)(m-2) - 2a. \tag{12}$$

Then

$$t + \left(\frac{r-g}{a}\right)'' + \sum_{i=4}^{m+1} Q_i \leqq \left(\frac{p_3 - c}{a}\right)'' - [4n/a] - 2n.$$

Multiply both sides of the inequality by a, substitute $aZ - aT$ for at, and then replace T by (12). This gives

$$\begin{aligned} aZ - 4an - a^2(2n+1) + ad + 2a + a(a-1)(m-2) \\ + 2a^2 + a\left(\frac{r-g}{a}\right)'' + a\sum_{i=4}^{m+1} Q_i \leqq a\left(\frac{p_3-c}{a}\right)'' - a[4n/a] - 2an. \end{aligned}$$

Increase $\left(\dfrac{r-g}{a}\right)''$ to $(r-g)/a(1+b) + b$ and $[4n/a]$ to $4n/a$. Then

$$\begin{aligned} \frac{r-g}{1+b} \leqq & -aZ + 4an + a^2(2n+1) - ad - 2a - 2a^2 \\ & - a(a-1)(m-2) - ab - a\sum_{i=4}^{m+1} Q_i + (1 + b \cdot 2^n)(1+b) \\ & - 4n - 2an - c/(1+b). \end{aligned} \tag{13}$$

Write $E = aZ - 4an - a^2(2n+1) + 2a + 2a^2 + a(a-1)(m-2) + ab + a\sum_{i=4}^{m+1} Q_i + c/(1+b) + 4n + 2an$. In (13) decrease g to 1. Then

$$(14) \qquad (r-1)/(1+b) \leqq -E - ad + (1 + b \cdot 2^n)/(1+b).$$

Similarly, for the second case from (11) let

$$(15) \qquad T = 4n + a(2n+1) - 2 - (a-1)(m-2) - 2a.$$

Proceeding as above, one obtains

$$(16) \qquad (r+1)/(1+b) \geqq E - q + d.$$

Let $m = 2n$. Then for $L_0 = p_{m+1} = 1 + b(2n)^n$, $L_0 v^n > ab$. For, when $n \geqq 19$, $(1 + b(2n)^n)v^n$ is approximately equal to $(1 + b(2n)^n)/n^n > 2^n b$. Now $2^n b > 4nb \geqq ab$ when $n \geqq 5$. It follows from this that $L_t V^n > b$. The conditions in Lemmas 12 and 13 are therefore satisfied.

For T in (12) and (15) it will now be shown that $t = Z - T \geqq 0$. For this purpose Z will be determined such that $\log_{10} L_{t+T} > \log_{10} N_1$. It suffices to take $\log_{10} N_1 = n^6$. In (8) take $v = 1/n$, decrease $L_0 = 1 + b(2n)^n$ to $b(2n)^n$ and increase a to $4n$, thereby increasing $t + T$. Then

$$\left(\frac{n}{n-1}\right)^{t+T} (n \log_{10} 2 - \log_{10} 4n) > n^6,$$

or

$$(17) \qquad Z = t + T > \frac{6 \log_{10} n - \log_{10} (n \log_{10} 2 - \log_{10} 4n)}{\log_{10} n - \log_{10} (n-1)}.$$

From (12) and (15) it is seen that $T > 0$ and $< 10n$. The numerator in (17) is $> 5 \log_{10} n$. It is therefore sufficient to show that

$$t + T > \frac{5 \log_{10} n}{\log_{10} n - \log_{10} (n-1)} > 10n.$$

By computation this inequality may be shown to be satisfied for $19 \leqq n \leqq 400$.

When $19 \leqq n \leqq 400$ the two conditions (14) and (16) are satisfied. Increase E to

$$(18) \qquad E_1 = 4nZ + 22n - 1 + a \sum_{i=4}^{2n+1} Q_i + 8n^2.$$

In (14) increase $r - 1$ to r and a to $4n$, multiply both sides by $1 + b$ and divide by $1 + 2^n b$. Then (14) holds if

$$r/(1 + 2^n b) \leqq 1 - (E_1 + 16n^2)(1+b)/(1 + 2^n b).$$

Similarly, (16) holds if

$$r/(1+2^nb) \geqq (E_1+4n)(1+b)/(1+2^nb).$$

Now

$$(1+b)/(1+2^nb) \leqq 2/(1+2^n).$$

The two inequalities above will therefore hold if

$$\frac{2E_1+8n}{1+2^n} \leqq \frac{r}{1+2^nb} \leqq 1-\frac{2(E_1+16n^2)}{1+2^n}. \tag{19}$$

The decimal part of $(1+3^nb)/(1+2^nb)$ is $r/(1+2^nb)$;

$$(1+3^nb)/(1+2^nb) = (3/2)^n - ((3/2)^n-1)/(1+2^nb).$$

The latter fraction decreases as n and b increase. When $n=19$ it is at most $.004^+$, $n=20$ at most $.003^+$. Hence the decimal part of $(1+3^nb)/(1+2^nb)$ agrees with that of $(3/2)^n$ to an increasing number of decimal places.

Let $n=19$. From (17), Z may be taken to be 302. From Lemmas 3 and 14,

$$a\sum_{i=4}^{2n+1} Q_i \leqq \sum_{i=4}^{2n+1} G_i = [p_4/p_3] + \cdots + [p_{2n+1}/p_{2n}] \leqq [(3/2)^n] + [(4/3)^n]$$
$$+ \cdots + [(2n/2n-1)^n] = [(3/2)^n] + \sum_{i=4}^{2n} K_i < [(3/2)^n] + X;$$

$X < 634.645589$. Also, $[(3/2)^{19}] \leqq 2217$. Hence, $E_1 \leqq 28708.645589$. From (19), $.111330 \leqq r/(1+2^nb) \leqq .866925$. Since $1-2(E_1+16n^2)/(1+2^n)$ increases with n and $(2E_1+8n)/(1+2^n)$ decreases when n increases, Theorem 3 will be true for any $n \geqq 19$ for which $r/(1+2^nb)$ lies between the above limits. The condition is satisfied for $n=19, 20$.

Let $n=20$. Then $Z=323$, $X<802.08$, $[(3/2)^{20}] \leqq 3326$, and $E_1 \leqq 33607.08$. Hence, $.064253 \leqq r/(1+2^nb) \leqq .923692$. This holds for $n=20, 21, \cdots, 28$.

Let $n=29$. Now $Z=522$, $X<7520.92208$, $[(3/2)^{29}] \leqq 127840$, and $E_1 \leqq 203277.92208$. Hence $.000758 \leqq r/(1+2^nb) \leqq .9991925$. This is true for $29 \leqq n \leqq 400$.

THE UNIVERSITY OF CHICAGO.

ON THE UNITS OF INDEFINITE QUATERNION ALGEBRAS.*

By Ralph Hull.

1. Introduction. The unit groups of maximal orders (integral sets) of definite quaternion algebras are finite groups which are easily determined for any given maximal order. In the indefinite case the unit groups are infinite and very little has been known of their structure until the recent work of Eichler,[1] whose results may be regarded as complete for those groups which contain no element $\neq \pm 1$ of finite order. In case the group contains nontrivial elements of finite order, Eichler determines the structure of an invariant subgroup, of finite index, free of such elements, from which the whole group is afterwards obtained by adjoining a suitable unit. It is the purpose of the present paper to present a method which applies equally well to both cases, and which has the added advantage, not possessed by Eichler's method, of providing the generators of the groups when desired.

The method will be indicated for maximal orders associated with canonical generations [2] of the algebras. It consists in associating the unit group with the norm form of the maximal order, written as a binary Hermitian form over an imaginary quadratic field. The units correspond to automorphs of this form and thus give rise to a principal circle group of transformations of the complex plane, for which, by methods due to Humbert,[3] a fundamental polygon may be constructed, and generators found. These are exhibited for some examples in the last section.

It is found that the principal circle groups involved are of finite character [4] $\{h, n\}$, where h is the genus of the surface associated with the fundamental polygon, and n is the number of its cycles of elliptic vertices. The class number n is evaluated by means of theorems of Kořínek [5] on relations between the ideals of maximal orders and those of splitting fields of the

* Received April 18, 1938.

[1] Eichler, *Mathematische Annalen*, vol. 114 (1937), pp. 637-654, cited hereafter as (E).

[2] These are due to Albert, *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 164-176, cited hereafter as (A).

[3] Humbert, *Comptes Rendus*, vol. 169 (1919), pp. 205-211.

[4] Fricke-Klein, *Automorphe Funktionen*, Chapter 3, cited hereafter as (F-K).

[5] Kořínek, *Mémoires de la Société Royale des Sciences de Bohême*, No. 1, 1932. For references to similar results of Hasse, Chevalley and E. Noether, see Deuring, "Algebren," *Ergebnisse der Mathematik* (1935), cited hereafter as (D).

algebras. Eichler's results give the non-Euclidean area of the fundamental polygon, and this, with general formulas of Fricke and Klein, subsequently gives the genus h.

2. The associated Hermitian form. Let $\mathfrak{Q}$ be a rational indefinite quaternion algebra with fundamental number $\sigma > 1$. Then σ is a product

$$(1) \qquad \sigma = q_1 q_2 \cdots q_{2r}$$

of an even number of distinct rational primes, and $\mathfrak{Q}$ is a division algebra of discriminant $-\sigma^2$. A canonical generation of $\mathfrak{Q}$ is of the form $\mathfrak{Q} = [1, i, j, ij]$, $i^2 = -p$, $j^2 = \sigma$, where p is a prime such that

$$(2) \qquad (q_s|p) = -1 \ (s = 1, \cdots, 2r), \qquad (\sigma|p) = 1.$$

We also write $\mathfrak{Q} = (\sigma, \mathfrak{Z})$, where $\mathfrak{Z}$ is the quadratic field $R(i)$, $i^2 = -p$.

The congruence

$$(4) \qquad 4\mu^2 \equiv \sigma \pmod{p}, \qquad 4\mu^2 - \sigma = kp,$$

has an integral solution μ which gives rise to a maximal order [6]

$$(5) \qquad \mathfrak{M} = [1, \omega, J, J\omega], \ \omega = (1 + i)/2, \ J = (2\mu + j)i/p.$$

The "similar" maximal order: $i^{-1}\mathfrak{M}i$ corresponds to the solution $-\mu$ of (4). We fix μ and write the general element v of $\mathfrak{M}$ in the form

$$(6) \qquad v = \xi + J\eta, \quad \xi = x_0 + x_1\omega, \quad \eta = y_0 + y_1\omega,$$

where the x and y are rational integers, and ξ, η are integral elements of $\mathfrak{Z}$.

The multiplication table of the basis (5) is readily computed and, in particular, we have the matrix representation of $\mathfrak{Q}$ corresponding to this basis:

$$(7) \qquad v = \xi + J\eta \rightarrow \begin{Vmatrix} \xi, & -2\mu x_1 - k\eta' \\ \eta, & \xi' - 2\mu y_1 \end{Vmatrix},$$

where ξ' denotes the conjugate of ξ in $\mathfrak{Z}$. The norm $N(v)$ is the determinant of the matrix (7) which may be written as the Hermitian form

$$(8) \qquad N(v) = N(\xi + J\eta) = \zeta\zeta' - \sigma\eta\eta'/p, \qquad \zeta = \xi + 2\mu i\eta/p,$$

and represents only rational integers for ξ, η integral in $\mathfrak{Z}$. The units of $\mathfrak{M}$ are its elements of norm ± 1. These evidently form a group which will be denoted by $\mathfrak{H}$.

---

[6] We use a slight modification of the form in (A).

Eichler has shown [7] that $\mathfrak{H}$ contains elements of norm -1 for every σ, and it is clear that the elements of norm $+1$ form an invariant subgroup $\mathfrak{H}_+$, of index 2 in $\mathfrak{H}$, and such that

$$\mathfrak{H} = \{u_1, \mathfrak{H}_+\}, \qquad N(u_1) = -1, \tag{9}$$

where u_1 is an arbitrary unit of norm -1. Henceforth we restrict attention to $\mathfrak{H}_+$.

Let u be an arbitrary unit in $\mathfrak{H}_+$:

$$u = \alpha + J\beta, \qquad N(u) = 1, \qquad \alpha \text{ and } \beta \text{ integral in } \mathfrak{Z}. \tag{10}$$

It follows readily from (7) that $uv = (\alpha + J\beta)(\xi + J\eta) = \xi_1 + J\eta_1$, where, on writing in accord with (8):

$$\gamma = \alpha + 2\mu i\beta/p, \qquad \zeta_1 = \xi_1 + 2\mu i\eta_1/p, \tag{11}$$

we have

$$\begin{aligned} \zeta_1 &= \gamma\zeta + \sigma\beta'\eta/p, \\ \eta_1 &= \beta\zeta + \gamma'\eta, \qquad \gamma\gamma' - \sigma\beta\beta'/p = 1. \end{aligned} \tag{12}$$

The form (8) is evidently invariant under (12). Although neither is integral in $\mathfrak{Z}$ we shall call the latter an *automorph* of the former and have

THEOREM 1. *The group H_+, can be represented as a group of automorphs* (12) *of the Hermitian form* (8).

3. The principal circle group. We now consider the transformation

$$T(z;u): \quad z_1 = \frac{\gamma z + \sigma\beta'/p}{\beta z + \gamma'}, \qquad \gamma\gamma' - \sigma\beta\beta'/p = 1, \tag{13}$$

of the complex z-plane, corresponding to the unit (10). Direct computation gives

$$z_1 z'_1 - \sigma/p = (zz' - \sigma/p)/|\beta z + \gamma'|^2, \tag{14}$$

where $|\beta z + \gamma'|$ denotes the absolute value. We see that $|\beta z + \gamma'| \neq 0$ for z interior to or on the circle

$$\mathfrak{C}: \quad zz' = \sigma/p. \tag{15}$$

For, if $\beta = 0$, $|\beta z + \gamma'| = \gamma\gamma' > 0$, and if $\beta \neq 0$, $|\beta z + \gamma'| = 0$ implies $z = -\gamma'/\beta$, $zz' = (p + \sigma\beta\beta')/p\beta\beta' > \sigma/p$ by (12). Hence $T(z;u)$ transforms $\mathfrak{C}$ into itself, its interior into its interior.[8]

---

[7] (E), p. 644.

[8] Units of norm -1 carry an interior into an exterior point.

We denote by $\mathfrak{U}$ the totality of transformations (13) corresponding to elements of $\mathfrak{H}_+$. Then $\mathfrak{U}$ is a group and we shall prove

THEOREM 2. *The group $\mathfrak{H}_+$ is homomorphic, in a two-to-one correspondence, with the group $\mathfrak{U}$, which is a Fuchsian group having the principal circle* (15).

The first part of the theorem is obvious since $T(z;u) = T(z;-u)$ and $N(-u) = N(u)$, and $N(u) = 1$ implies that α and β have g. c. d. 1 in $\mathfrak{Z}$. To prove the last part we have only to show [9] that $\mathfrak{U}$ is properly discontinuous on the interior of $\mathfrak{C}$. At the same time we give formulas which are useful for finding units.

We write

$$\begin{aligned}
&\alpha = a_0 + a_1\omega, \qquad \beta = b_0 + b_1\omega,\\
&X = i(\gamma' - \gamma) = pa_1 + 2\mu(2b_0 + b_1),\\
(16)\qquad &Y = \beta + \beta' = 2b_0 + b_1,\\
&Z = -i(\beta - \beta')/p = b_1,\\
&W = \gamma + \gamma' = 2a_0 + a_1 - 2\mu b_1.
\end{aligned}$$

Then $X, \cdots, W$ are rational integers such that

$$\begin{aligned}
&X^2 + pW^2 - \sigma(Y^2 + pZ^2) = 4p,\\
(17)\qquad &X \equiv 2\mu Y \pmod p), \qquad X \equiv W, \qquad Y \equiv Z \pmod 2),\\
&4N(\beta) = Y^2 + pZ^2, \qquad 4pN(\gamma) = X^2 + pW^2.
\end{aligned}$$

It is evident from (17) that at most a finite number of units correspond to a given β. We have also

$$\begin{aligned}
(18)\qquad |\beta z + \gamma'| &= |\gamma'| \cdot |1 + \beta z/\gamma'|\\
&> |\gamma'| \cdot (1 - |\beta z/\gamma'|)\\
&> |\gamma'| \cdot (1 - \beta\beta' zz'/\gamma\gamma')/(1 + |\beta z/\gamma'|).
\end{aligned}$$

For $zz' < \sigma/p$, it follows from (17) and (18) that $|\beta z + \gamma'| \to \infty$ as $|\beta| \to \infty$ since the last fraction in (18) has a finite limit $\neq 0$. Hence, by (14), the transforms of an arbitrary point interior to $\mathfrak{C}$ approach the circumference of $\mathfrak{C}$ as $N(\beta) \to \infty$. It follows that a cluster point of such transforms is necessarily on the circumference [10] and $\mathfrak{U}$ is properly discontinuous on the interior of $\mathfrak{C}$. Hence $\mathfrak{U}$ is a Fuchsian group with principal circle $\mathfrak{C}$ by the definition of such groups.

---

[9] Our proof is essentially that of Picard, *Annales Scientifiques de l'École Normale Supérieure*, vol. 20 (1884), p. 50.

[10] That, conversely, every point of the circumference is a cluster point of transforms of an interior point follows indirectly from later results, no. 5.

4. The fundamental polygon. We wish to apply Humbert's practical formulation of the method of "rayonnement" to find a fundamental domain for $\mathfrak{U}$ interior to $\mathfrak{C}$. For this purpose, it is necessary to assume that $z=\infty$ is not a fixed point of a substitution $T(z;u)$, $u \neq \pm 1$. This is evidently equivalent to the assumption

$$u \neq \pm 1 \textit{ implies } \beta \neq 0. \tag{19}$$

If $p=3$, we have the unit $u=\alpha=(-1+i)/2$, $\beta=0$. *If* $p>3$, $\alpha\alpha'=1$ *has the only integral solutions* $\alpha=\pm 1$. We *assume*[11] *henceforth that* $p>3$. Then we have (19) and[12]

THEOREM 3. *If $p>3$, the group $\mathfrak{U}$ has a fundamental domain, for the interior of $\mathfrak{C}$, consisting of that part of the interior of $\mathfrak{C}$ which is exterior to all the circles*

$$|\beta z+\gamma'|^2=1, \tag{20}$$

as u ranges over $\mathfrak{H}_+$.

The circles (20) are orthogonal to $\mathfrak{C}$ and are called[13] the isometric circles of the substitutions of $\mathfrak{U}$. The boundary of the fundamental domain is a polygon whose sides are arcs of certain of the circles (20). We denote this fundamental polygon by $\mathfrak{P}$.

5. Properties of the fundamental polygon. We require certain formulas, related to the polygon $\mathfrak{P}$, which are obtained by combining the results of Eichler[14] with well-known formulas from the theory of Fuchsian groups and the associated hyperbolic geometry.

The interior of $\mathfrak{C}$, with its orthogonal circles as "straight lines," is the hyperbolic plane of the geometry of Lobatchefski. For this geometry, $\mathfrak{U}$ is a group of displacements preserving the projective metric. Eichler, in a somewhat less explicit way than that employed in the present paper, represented $\mathfrak{H}$ as a group of non-Euclidean displacements of the interior of the *unit* circle $\mathfrak{C}_1$. He found, by analytic methods, the formula

$$\mathfrak{A}=\pi\phi(\sigma)/3, \tag{21}$$

---

[11] This is a restriction only when $R(i)$, $i^2=-3$, splits $\mathfrak{Q}$. For any σ, there exist infinitely many p giving canonical generations.

[12] Humbert, *loc. cit.*, 3).

[13] Ford, *Automorphic Functions*, p. 23.

[14] (E), p. 649. Cf. Humbert, *Comptes Rendus*, vol. 171 (1920), pp. 377-383, for somewhat similar results obtained by the same methods directly for Hermitian forms. For the general theory of the polygons associated with Fuchsian groups, see (F-K), in particular: p. 262.

where ϕ is the Euler function, for the non-Euclidean area $\mathfrak{A}$ of a fundamental domain interior to $\mathfrak{C}_1$. We obtain $\mathfrak{C}_1$ from $\mathfrak{C}$ by the linear transformation

$$z_1 = (p/\sigma)^{\frac{1}{2}}z,$$

which preserves the projective metric. Moreover, the non-Euclidean area of a fundamental domain for $\mathfrak{U}$ interior to $\mathfrak{C}$ is independent of the particular choice of the domain since $\mathfrak{U}$ also preserves the metric. It follows that (21) gives the area $\mathfrak{A}$ of the polygon $\mathfrak{P}$.

On the other hand, the projective area of a polygon interior to $\mathfrak{C}$ is given by the formula

$$\mathfrak{A} = (m-2)\pi - \Sigma, \tag{22}$$

where m is the number of sides of the polygon (arcs of circles orthogonal to $\mathfrak{C}$), and Σ is the sum of its angles (measured in the elementary sense). We combine (21) and (22) to obtain

THEOREM 4. *Let m be the number of sides, Σ the sum of the angles, of $\mathfrak{P}$. Then*

$$\pi\phi(\sigma)/3 = (m-2)\pi - \Sigma. \tag{23}$$

The number of sides of $\mathfrak{P}$ is even: $m = 2t$, with the understanding that two sides, forming the angle π, may lie along the same arc in case a "non-apparent" vertex of $\mathfrak{P}$ occurs. This corresponds to a fixed point of an elliptic substitution of order 2. We proceed to a further study of the vertices of $\mathfrak{P}$.

The finiteness of $\mathfrak{A}$ implies that $\mathfrak{P}$ has at most parabolic vertices on $\mathfrak{C}$, and we now show that such do not actually occur. For, a parabolic vertex of $\mathfrak{P}$ is the fixed point of a substitution $T(z;u)$, $u \neq \pm 1$, with $\gamma + \gamma' = 0$. By (13) such a fixed point is of the form $z = (\gamma - \gamma')/2\beta$, and $zz' = \sigma/p$ implies $N(\gamma - \gamma' + 2J\beta) = 0$. This is impossible since $\mathfrak{Q}$ is a division algebra.[15] It follows that the cycles of vertices of $\mathfrak{P}$ are of at most two types: (1) adventive (*zufällig*) vertices, for which the sum of the angles is $2\pi = 2\pi/k$, $k = 1$; (2) elliptic vertices, for which the sum of the angles of a cycle is $2\pi/k$, $k > 1$, where k is the order of the elliptic substitutions of the corresponding class. In the next section it will be seen that type (2) may or may not occur and that, for this type, k can have only the values 2 or 3. We denote by m_1, m_2, m_3 the number of cycles of vertices of $\mathfrak{P}$ for which $k = 1, 2, 3$, respectively. Evaluating Σ, and inserting the result in (23), we obtain

$$\phi(\sigma) = 6(t - m_1) - 6 - 3m_2 - 2m_3. \tag{24}$$

---

[15] This property of $\mathfrak{Q}$ is also essential in Eichler's proof of (21).

By the usual identification of equivalent points on the sides of $\mathfrak{P}$ a surface is obtained whose genus h is given by the formula [16]

$$2h - 1 = t - m_1 - m_2 - m_3. \tag{25}$$

We combine (24) and (25) to obtain (26) below and have

THEOREM 5. *The polygon $\mathfrak{P}$ is associated with a surface of genus h given by*

$$12(h-1) = \phi(\sigma) - 3m_2 - 4m_3, \tag{26}$$

where m_2 and m_3 are the number of cycles of elliptic vertices of $\mathfrak{P}$ corresponding to classes of elliptic substitutions of orders 2 and 3 respectively.

6. The classes of elliptic substitutions. We recall that two elliptic substitutions of a Fuchsian group are said to be in the same class if one is the transform of the other by a substitution of the group. For the group $\mathfrak{U}$, the numbers m_2 and m_3, defined in Theorem 5, have the values given in

THEOREM 6. *The group $\mathfrak{U}$ has elliptic substitutions at most of orders $k = 2$ and $k = 3$. The case $k = 2$ occurs, that is $m_2 > 0$, if and only if every odd prime divisor of σ is $\equiv 3 \pmod 4$ and then $m_2 = 2^{2r}$ or $m_2 = 2^{2r-1}$, according as $2 \nmid \sigma$ or $2 \mid \sigma$. The case $k = 3$ occurs, that is $m_3 > 0$, if and only if every prime divisor $\neq 3$ of σ is $\equiv 2 \pmod 3$ and then $m_2 = 2^{2r}$ or $m_3 = 2^{2r-1}$, according as $3 \nmid \sigma$ or $3 \mid \sigma$. In both cases, r is defined in* (1).

An elliptic substitution $T(z;u)$ of $\mathfrak{U}$ corresponds to a unit $u \neq \pm 1$ of $\mathfrak{H}_+$ which is of finite order and hence a root of unity. Since $\mathfrak{Q}$ is of degree 2 over R, $\mathfrak{Q}$ can contain at most fourth or sixth roots of unity. This proves the first sentence of the theorem since $T(z;u) = T(z;-u)$, and in case $k = 3$, one of the units $\pm u$ is a cube root of unity. In view of the last remark it will be sufficient to give details only for the case $k = 2$ since those for $k = 3$ are exactly similar.

By known theorems [17] on the splitting fields of algebras, a quadratic field $\mathfrak{F}$ splits $\mathfrak{Q}$ if and only if no prime divisor of σ is the product in $\mathfrak{F}$ of two distinct prime ideals. This gives at once the criterion of the theorem for the case $k = 2$ (and similarly for $k = 3$) by the known laws of factorization in $\mathfrak{F} = R(\theta)$, $\theta^2 = -1$. We assume henceforth that the criterion is fulfilled and proceed to evaluate m_2.

---

[16] (F-K), p. 262, formula (2).

[17] General theorems are due to Hasse, Cf. (D), pp. 117-118. For the present special case of quaternion algebras elementary proofs are given by Latimer, *Duke Mathematical Journal*, vol. 2 (1936), pp. 681-684.

First, let u and u' be any two units of $\mathfrak{H}_+$ such that $T(z;u)$ and $T(z;u')$ are elliptic of order 2. Then $\mathfrak{F}_1 = R(u)$ and $\mathfrak{F}_2 = R(u')$ are quadratic sub-fields of $\mathfrak{Q}$ isomorphic under the correspondence $u \leftrightarrow u'$. Hence [18] there exists a regular element v of $\mathfrak{Q}$ such that $v^{-1}uv = u'$. Write $\mathfrak{M}_1 = v\mathfrak{M}v^{-1}$, and let $\mathfrak{m}$ denote the unique maximal order of $\mathfrak{F}_1$. Without loss of generality, we may assume that v is in $\mathfrak{M}$. Then $v\mathfrak{M}$ is an $\mathfrak{M}$-right-ideal whose left order $\mathfrak{M}_1$ contains $\mathfrak{m} = v\mathfrak{m}'v^{-1}$, where $\mathfrak{m}'$ is the maximal order of $\mathfrak{F}_2$. Hence [19] the *distance ideal* $\mathfrak{D} = (\mathfrak{M} \cdot \mathfrak{M}_1)^{-1} = \mathfrak{a}\mathfrak{M}$, where $\mathfrak{a}$ is an $\mathfrak{m}$-ideal. We have also $\mathfrak{M}_1 = \mathfrak{D}\mathfrak{M}\mathfrak{D}^{-1}$. Since $R(\theta)$, $\theta^2 = -1$, has class number 1, $\mathfrak{a}$ is a principal ideal: $\mathfrak{a} = \alpha\mathfrak{m}$, α in $\mathfrak{m}$, and $\mathfrak{M}_1 = \alpha\mathfrak{M}\alpha^{-1}$. A rational number a can be chosen so that $a\alpha^{-1}v = v_1$ is a primitive [20] element in $\mathfrak{M}$. Then $v_1^{-1}uv_1 = u'$, $v_1^{-1}\mathfrak{M}v = v^{-1}\alpha\mathfrak{M}\alpha^{-1}v = v^{-1}\mathfrak{M}_1 v = \mathfrak{M}$, and $v_1\mathfrak{M} = \mathfrak{M}v_1 = \mathfrak{J}$ is a primitive two-sided ideal of $\mathfrak{M}$. This proves the

LEMMA. *Let u and u' be any two imaginary fourth roots of unity in $\mathfrak{M}$. Then there exists a two-sided primitive integral ideal $\mathfrak{J}$ of $\mathfrak{M}$ such that $\mathfrak{J}$ is a principal ideal and*

$$\mathfrak{J} = v\mathfrak{M} = \mathfrak{M}v, \qquad v^{-1}uv = u'. \tag{27}$$

In view of the Lemma, to complete the proof of Theorem 6 we have to count the two-sided primitive ideals $\mathfrak{J}$ of $\mathfrak{M}$, and then determine when the substitutions $T(z;u)$ and $T(z;u')$, corresponding to u and u' in (27), are in the same class.

The primitive two-sided ideals [21] of $\mathfrak{M}$ are the 2^{2r} products

$$\mathfrak{J}(\epsilon) = \mathfrak{q}_1^{\epsilon_1} \cdots \mathfrak{q}_{2r}^{\epsilon_{2r}}, \qquad \epsilon_s = 0, 1 \ (s = 1, \cdots, 2r), \tag{28}$$

where $\mathfrak{M}\mathfrak{q}_s = \mathfrak{q}_s\mathfrak{M} = \mathfrak{q}_s$, $\mathfrak{q}_s^2 = q_s$, for the factors (1) of σ. Since [22] the class number of $\mathfrak{Q}$ is 1, each $\mathfrak{J}$ in (28) is principal: $\mathfrak{J}(\epsilon) = v(\epsilon)\mathfrak{M} = \mathfrak{M}v(\epsilon)$.

Let $\mathfrak{J} = v\mathfrak{M}$ be a fixed ideal (28) and suppose that $T(z;u)$ and $T(z;u')$, $u' = v^{-1}uv$, are in the same class. Then there exists a unit ξ of $\mathfrak{H}_+$ such that $u' = \xi^{-1}u\xi$. Since $\xi v^{-1}uv\xi^{-1} = \xi u'\xi^{-1} = u$, and $R(u)$ is a maximal sub-field of $\mathfrak{Q}$, it follows that $v\xi^{-1}$ is in $R(u)$. But $v\xi^{-1}$ is integral and $N(v\xi^{-1}) = N(v)N(\xi^{-1}) = N(v) = q_1^{\epsilon_1} \cdots q_{2r}^{\epsilon_{2r}}$, from which it follows that at most the exponent ϵ corresponding to $q = 2$ can be $\neq 0$ since the remaining q's are indecomposable in $R(u)$. Thus, if $2 \nmid \sigma$, $N(v) = 1$, v is a unit, and $m_2 = 2^{2r}$. If $2 \mid \sigma$, the right-ideal $\mathfrak{J}_1 = (1+u)\mathfrak{M}$, $N(\mathfrak{J}_1) = 2$, is necessarily two-sided

[18] (D), p. 42, Theorem 3.

[19] Kořínek, *loc. cit.*, 5), p. 9, Theorem 3.

[20] That is, not divisible by a rational integer > 1.

[21] See (D), p. 88, Theorem 2.

[22] Eichler, *Journal für Mathematik*, vol. 176 (1937), p. 192.

since 2 has only two-sided ideal factors in $\mathfrak{M}$, and hence $(1+u)\mathfrak{M}=v\mathfrak{M}$, $v=(1+u)\eta$, $u'=v^{-1}uv=\eta^{-1}(1+u)^{-1}u(1+u)\eta=\eta^{-1}u\eta$, where η is a unit of $\mathfrak{H}_+$. In this case evidently $m_2=2^{2r-1}$. This completes the proof of Theorem 6 for $k=2$.

7. The structure of the unit groups. We have seen that $\mathfrak{H}$ is of the form (9) and

$$\mathfrak{H}_+=\{-1,\mathfrak{U}\}, \tag{29}$$

where $\mathfrak{U}$ is a Fuchsian group of finite character $\{h,n\}$, $n=m_2+m_3$, $12(h-1)=\phi(\sigma)-3m_2-4m_3$, and m_2 and m_3 are given by Theorem 6. An immediate consequence of the general theory[23] of Fuchsian groups is

THEOREM 7. *The group* $\mathfrak{U}$ *has a canonical set of generators:*

$$\mathfrak{U}=\{u_1,\cdots,u_n;\ v_1,\cdots,v_h;\ v'_1,\cdots,v'_h\}, \tag{30}$$

which satisfy the only essential relations

$$\prod_{i=1}^{n}u_i\prod_{j=1}^{h}v_j^{-1}v'_jv_jv'^{-1}_j=1,\qquad u_i^k=1,\qquad k=2 \text{ or } 3. \tag{31}$$

The group $\mathfrak{H}$ *is of the form* (9), (29), (30) *and* (31).

In case $m_2=m_3=n=0$, that is, by Theorem 6, if at least one prime divisor of σ is $\equiv 1 \pmod 4$ and at least one $\equiv 1 \pmod 3$, the generators u in (30), (31) are lacking and the form of Theorem 7 coincides with that given by Eichler.[24] In case $n\neq 0$, that is at least one of m_2 and $m_3\neq 0$, Eichler studied a congruentially defined sub-group $\mathfrak{H}_\mathfrak{q}$ of $\mathfrak{H}:\mathfrak{H}_\mathfrak{q}\equiv 1 \pmod{\mathfrak{q}}$, where $\mathfrak{q}$ is a suitable $\mathfrak{q}$ in (28). The corresponding $\mathfrak{U}_\mathfrak{q}$ is of character $\{h_\mathfrak{q},0\}$, $24(h_\mathfrak{q}-1)=\phi(\sigma)(q+1)$, $N(\mathfrak{q})=q$, and

$$\mathfrak{H}=\{\eta,\mathfrak{U}_\mathfrak{q}\},\qquad \eta^{2(q+1)} \text{ in } \mathfrak{U}_\mathfrak{q},\qquad \eta^{-1}\mathfrak{U}_\mathfrak{q}\eta=\mathfrak{U}_\mathfrak{q},$$

where $\mathfrak{U}_\mathfrak{q}$ has generators v, v' as in (30) and (31), with h replaced by $h_\mathfrak{q}$.

In general, the polygon $\mathfrak{P}$ of **4** does not give the generators (30) immediately. They may subsequently be found by the usual transformations from the generators associated with $\mathfrak{P}$ and the relations satisfied by them.

8. Examples. For a few small values of σ we give some results of actual computation of generators of the groups $\mathfrak{U}$. These are determined for canonical maximal orders $\mathfrak{M}$, but the unit groups of all maximal orders of a given $\mathfrak{Q}$ are isomorphic since $\mathfrak{Q}$ has class number 1.

---

[23] (F-K), pp. 186-187.

[24] (E), p. 650.

The units are obtained from solutions (X, Y, Z, W) of (16) and (17) and it will be convenient to speak of the unit (X, Y, Z, W) and of the (isometric) circle (X, Y, Z, W). We note that the units (X, Y, Z, W) and $(X, Y, -Z, -W)$ are inverses and one carries its own circle into that of the other. Also, the circles (X, Y, Z, W) and $(X, Y, -Z, W)$ are symmetric with respect to the axis of imaginaries. The cases $k = 2$ and $k = 3$ of Theorem 6 correspond to $W = 0$ and $W = \pm 1$, respectively. By analogy with the Pell equation, it is to be expected that the sides of $\mathfrak{P}$ will correspond to small values of $4N(\beta) = Y^2 + pZ^2$. However, certain units may be found, before $\mathfrak{P}$ is closed, which do not contribute to the boundary of $\mathfrak{P}$. Such units are contained in the subgroups generated by units found earlier. In some cases, e. g. $\sigma = 10$, this increases the computations. The examples follow.

I. $\sigma = 6,\ p = 19,\ \mu = -7,\ \{h, n\} = \{0, 4\},\ m_2 = m_3 = 2.$

$\mathfrak{P}$ is bounded by the five circles $(10, 2, 0, 0)$, $(1, 4, 0, \pm 3)$ and $(-14, 1, \pm 1, 0)$. The first contributes two sides since the point at which it crosses the axis of imaginaries is a fixed point of the corresponding $T(z; u)$. In this case, $\mathfrak{U}$ can be generated by three of its substitutions (and their inverses). Canonical generators $u_1, \cdots, u_4$ are $(10, 2, 0, 0)$, $(15, 3, 1, -1)$, $(14, -1, 1, 0)$ and $(-9, 2, 0, 1)$, respectively.

II. $\sigma = 10,\ p = 43,\ \mu = 29,\ \{h, n\} = \{0, 4\},\ m_2 = 0,\ m_3 = 4.$

For $\sigma = 10$, $p = 3$ gives a canonical generation of Q but is excluded by **4**. $\mathfrak{P}$ is bounded by $(-13, 2, 0, \pm 1)$, $(17, 4, 0, \pm 1)$ and $(8, 12, 0, \pm 6)$. Canonical generators $u_1, \cdots, u_4$ are $(-13, 2, 0, 1)$, $(17, 4, 0, -1)$, $(47, 6, 2, -1)$ and $(43, 0, 2, -1)$, respectively.

III. In each of the cases $\sigma = 14 (p = 11)$ and $\sigma = 15 (p = 7)$, $\mathfrak{P}$ has ten sides, with a non-apparent vertex in case $\sigma = 14$.

IV. $\sigma = 26,\ p = 11,\ \mu = 1,\ \{h, n\} = \{2, 0\},\ m_2 = m_3 = 0.$

This is the smallest case of σ for which $n = 0$. $\mathfrak{P}$ has fourteen sides and its vertices are grouped in four cycles, two of three and two of four vertices each.

The writer has been unable to determine generally the number of sides of $\mathfrak{P}$. This, with the number m_1 of non-elliptic cycles of vertices, was eliminated from (24) and (25). Inasmuch as $\mathfrak{P}$ is a special *domaine rayonné*, with its "center" coinciding with the center of $\mathfrak{C}$, it is possible that, with a different choice of "center," a polygon more suitable for finding canonical generators (30) and (31) could be found.

THE UNIVERSITY OF ILLINOIS.

NON-ABELIAN COMPACT CONNECTED TRANSFORMATION GROUPS OF THREE-SPACE.*

By DEANE MONTGOMERY and LEO ZIPPIN.

1. In this paper we conclude our study of *compact connected effective transformation groups of Euclidean three-space* (hereafter, $Tg(E_3)$) with a final theorem on the non-abelian case. We have already established and shall use the result that:

1′) *The only abelian $Tg(E_3)$ is, abstractly, the group of rotations of a circle.*[1] *Moreover, if K denotes such a group, the space E_3 must admit a coördinate system in which K is the group of all rigid rotations of E_3 about a fixed axis.*[2]

This will now be complemented by the following theorem:

1″) *The only non-abelian $Tg(E_3)$ is, abstractly, the group of rotations of the two-sphere.*[3] *Moreover, if G' denotes such a group, the space E_3 must admit a coördinate system in which G' is the group of all rigid rotations of E_3 about the origin.*

The plan underlying our proof is to show that a group G', as in 1″, contains at least two distinct groups as in 1′ and that there exists a *single* coördinate system in E_3 which is appropriate to these two groups. Interlocked with this argument there is a topological analysis of the orbits of points under the group generated by these two, by which we achieve a reduction of our problem from three- to two-space.

1.1 We remind the reader that a topological group G is called a transformation group of a topological space R provided that with each element g of G and point x of R there is associated a unique point $g(x)$ of R and:

i) the association $g(x)$ is simultaneously continuous in g and x,
ii) when g_0 is the identity of G, $g_0(x) = x$,
iii) when $g_3 = g_1g_2$, $g_1[g_2(x)] = g_3(x)$.

---

* Received June 18, 1938.

[1] We shall call this the circle-group: it may also be realized as the additive group of real numbers modulo one.

[2] In cylindrical coördinates the fixed axis will be the line $r = 0$, and the elements of K will be the transformations: $z' = z$, $r' = r$, $\theta' = \theta + 2\pi t$ $(0 \leqq t < 1)$. See (4) and (5) of the bibliography at the end of the paper.

[3] The structure of this group is well-known.

The transformation group is said to be *effective* if

$$g(x) = x, \text{ for } all \ x$$

is characteristic of the identity element *only*.[4]

1.2 We recall the obvious consequences that each element of G is a homeomorphic mapping of R onto all of R and that for each x the "orbit" $G(x)$,[5] the set of all "transforms" of x under G, is homeomorphic to a coset-space of G (depending, in general, on x). Accordingly, if G is compact and connected, so is $G(x)$. Always, $G(x)$ is a "strongly-homogeneous" subset of R. This means, by definition, that if y and z denote two points of $G(x)$ there exists at least one homeomorphic mapping of R onto itself under which $G(x)$ is invariant and y is carried to z.

We remark that if G is an effective transformation group of R, all subgroups of G are likewise effective.

2. We now prove a general group-theoretic lemma.

LEMMA 1. *Every compact connected non-abelian group G' contains at least two distinct compact connected abelian subgroups K_1 and K_2.*

It is known (7) that such a group G' must contain (arbitrarily small compact abelian invariant) subgroups G^* for which G'/G^* is a non-abelian Lie group. The Lie group G/G^* must (see e. g. 2), contain a one-parameter subgroup. The closure of this we will denote by $\bar{F}$. The group $\bar{F}$ is an abelian Lie group and therefore a *toral* group (6, 7). Then $\bar{F}$ contains an element $\bar{f}$ whose powers are everywhere dense in $\bar{F}$. Let f be an element of G in the "coset" $\bar{f}$. The powers of f form a group whose closure F is abelian. The factor group of F by its intersection with G^* must "cover" $\bar{F}$. Then it is clear that F cannot be totally disconnected and must contain a connected subgroup K_1. Now G/G^* cannot coincide with $\bar{F}$ (which is abelian) and must therefore contain a one parameter group with elements not in $\bar{F}(2)$. From this second group we obtain, quite as above, an abelian group K_2 which must contain elements not in K_1.

3. For the remainder of this paper G' will denote a fixed non-abelian $Tg(E_3)$ and E will denote the euclidean three-space (heretofore E_3). The subgroups K_1 and K_2, as in the previous lemma, are themselves *abelian* $Tg(E)$

[4] Each transformation group G carries with it, in a natural way, an effective factor-group. See (5).

[5] In general, if K denotes a subgroup of G and S a subset of R, $K(S)$ will denote the set of all points $k(s)$ for k of K and s in S.

and therefore circle-groups. From the fact that they are distinct subgroups of G' we shall show that *there is at least one point x of E such that the orbits $K_1(x)$ and $K_2(x)$ are distinct point sets.* To do this we shall suppose the contrary—that for *every* x, $K_1(x) = K_2(x)$ and obtain a contradiction.

3.1 Let A denote the set of points each of which is fixed under K_1. This set is closed in E and homeomorphic to a line. Let x by any point in the complement of A. Then $K_1(x) = K_2(x)$ is a simple closed curve. Let g_1 denote an element of K_1 which is not in K_2. This element exists for, otherwise, K_1 would be a subgroup of K_2 and could not then be distinct from it since both are circle groups. Let y be the point $g_1(x)$. The point y is distinct from x because g_1 is not the identity and x is not in A. Since $y \subset K_2(x)$ there is an element g_2 of K_2 such that $g_2(y) = x$. Then $g_2 g_1$ is an element g of G' such that $g(x) = x$.

3.2 We shall now show that g is the identity. Let F denote the points of $E - A$ which are fixed points of g. This set is *closed in* the set $E - A$ and is not vacuous, since it contains x. If we can show that F is also *open in* $E - A$ we shall have accomplished our immediate object, for in this case, since $E - A$ is both open and connected, it will follow that F must coincide with $E - A$. On the other hand the fixed points under g form a closed set, necessarily, and this set must contain $\bar{F} = \overline{E - A} = E$. But if all points of E are fixed points of g, g *is* the identity.

3.3 Let us show, then, that F is open. To this end, let z denote an arbitrary point of F. By our assumption in 3, $K_1(z) = K_2(z)$. From this and the composition of g, $K_1(z)$ is invariant under g. This means that g may be regarded as a homeomorphic mapping of a *simple closed curve into itself*. But g has a fixed point, namely z. Furthermore g cannot reverse orientation on $K_1(z) = K_2(z)$ since this would imply that either g_1 or g_2 is sense-reversing. This is impossible because each of them is an element of a *connected* group (K_1, K_2 resp.). Finally, the further condition upon g that its powers must generate a *compact* group obliges it, as is well known, to be the identity as a transformation of $K_1(z)$. That is, $g(z') = z'$ for every point z' of $K_1(z)$. Then, of course, every point of $K_1(z)$ is fixed under all powers of g.

The last remark makes it clear that points sufficiently near to z are moved arbitrarily little by all powers of g and it is an immediate consequence of this that all orbits sufficiently close to $K_1(z)$ have on them at least one fixed point under g. But then, by the argument above, such orbits are pointwise fixed under g and, finally, the set F is open in $E - A$.

3.4 Now, at last, if g is the identity of G' it follows that g_1 is the inverse of g_2 and this contradicts its choice as an element not in K_2.

4. Throughout most of the paper we shall be concerned with the group, we denote it by G, which is the *closure* of the group generated by K_1 and K_2. It will transpire that G is identical with the original G'.

4.1 No orbit $G(x)$ can be three-dimensional, for if it were it would have to contain inner points (relative to E) and it would thereupon be open, since it it strongly homogeneous. But it is closed and so would coincide with E. This it cannot do, being compact. It follows that orbits $G(x)$ are at most two-dimensional. We shall learn, much later, that these orbits, save one, are in fact two-spheres. For the moment, it is clear that if y is a point not on some $G(x)$ there is at least one point z of $G(x)$ which is accessible[6] from y. It follows, from the strong homogeneity of $G(x)$, that every point of that set is accessible from at least one point of $G(y)$. Since $G(y)$ is connected we may conclude that every point of $G(x)$ is likewise accessible from y. This establishes:

4.1′) *$G(x)$ is the boundary of every complementary domain and each point of $G(x)$ is accessible from each domain.*

This implies, of course, that *no subset of $G(x)$ can separate E.*

4.2 We shall now introduce into E a (cylindrical[7]) coördinate system appropriate to the "rotation" character of K_1. In this (z, r, θ)-system the set of points which are fixed points under each transformation of K_1 constitute the axis $A: r = 0$. This axis is the edge of the closed half-plane P which we may take as the "initial plane" $\theta = 0$ of our coördinate system. Each orbit $K_1(x)$ intersects the half-plane P in precisely one point which we may call the initial point of the orbit.

4.2′) We shall denote by $\theta\{x\}$ the single-valued continuous function defined everywhere in E with point-values in P which maps each point into the initial point of its orbit.

It may assist the reader if he will observe that the orbits $K_1(x)$ are actual circles[8] in our coördinate system: if $\bar{x}$ is the point $(\bar{z}, \bar{r}, \bar{\theta})$ then $K_1(x)$ is the set $z = \bar{z}$, $r = \bar{r}$. The orbits $K_2(x)$ however are merely simple closed curves[8] (topological circles). While we may interchange the rôles of K_1 and K_2 if

---

[6] I. e. there exists an arc yz, $yz \cdot G(x) = z$.

[7] We choose this for definiteness; we shall make only nominal use of the form of the coördinate system and none, for the present, of the coördinates themselves.

[8] The fixed points excepted, of course.

we wish, we cannot, at this moment, find a coördinate system in which both sets of orbits are circles.

4.3 Since the orbit $G(x)$, for every point x, is invariant under K_1 it is clear that the set $P_x = P \cdot G(x)$ is identically the set $\theta\{G(x)\}$ and that $G(x) = K_1(P_x)$ is a "figure of rotation." We know that $G(x)$ is a homogeneous compact connected set. In the next section we shall prove that $G(x)$ is locally connected.

5. Let x denote an arbitrary point for which $K_2(x)$ is distinct from $K_1(x)$, section 3. We may suppose, without loss, that x is the initial point of its orbit, a point of P. Then $\theta\{K_2(x)\}$ will contain x. On the other hand, since $K_2(x)$ is not a subset of $K_1(x)$ and certainly not a subset of A, $K_2(x)$ cannot be a subset of $K_1(x) + A$ (whether or not x is a point of A) and consequently $\theta\{K_2(x)\}$ is not a subset of $x + A$. It follows, from the fact that $\theta\{K_2(x)\}$ is a Peano space, that it must contain some arc yz (for some pair of points y and z) which has no point in A. Now $K_1(yz)$ must contain a 2-cell which we will denote by S (4). The two-cell S, by its construction, must be a subset of the orbit $G(x)$.

5.1 Let s denote an inner point of the two-cell S. We propose to show, here, that $G(x)$ is locally connected at s. If this is not the case then there must exist an open set O containing s and a sequence of points s_n of $G(x)$ such that $\lim s_n = s$ and such that s and s_n are in no connected subset of $O \cdot G(x)$. There exists a sequence of elements g_n of G such that $s_n = g_n(s)$ and such that the g_n's converge to the identity.[9] Now let S' denote a 2-cell subset of S which has s as inner point and which is contained in O. Almost all of the two-cells $g_n(S')$ will lie in O and no two of these can have a point in common, by our choice of O. There is an arc ts which has s only on $G(x)$, by 4.1'. The two-cell S' must have two "sides" in E at the point s. Then it is clear that the arc ts must approach S' from *one* of these sides while the sets $g_n(S')$, which converge to S', must approach it from the other. Since $g_n(s)$ is a point of $g_n(S')$ it follows easily that for at least one large n, $g_n(ts)$ and S' have a common point distinct from s. But this is a contradiction of the fact that $g_n(ts)$ and $G(x)$ can have s_n, only, in common.

5.1' Then we have shown that for every point x for which $K_1(x) \neq K_2(x)$ the set $G(x)$ is locally connected, and is two-dimensional. Then the set

[9] Let g'_n denote a sequence such that $s_n = g'_n(s)$. We may take it, from the compactness of G, that there is a $g = \lim g'_n$. From the simultaneous continuity $g(s) = s$. Then $g^{-1}(s) = s$ and $g'_n g^{-1}(s) = s_n$. Let $g_n = g'_n g^{-1}$.

$P_x = P \cdot G(x)$ is locally connected, and we know that it is not a subset of $x + A$.

5.1″ Suppose, now, that y is a point for which $K_1(y) = K_2(y)$. From the fact that every g of G is the limit of finite products of elements from K_1 and K_2 it follows, at once, that $G(y) = K_1(y)$. From this it is clear that P_y is a point.

5.2 We have shown now that P_x is a Peano space and not a subset of A *or* that it is a point. We have seen, also, that in the first alternative at least one (and therefore every) point is interior to some two-cell of $G(x)$.[10] It should be reasonably clear that $G(x) = K_1(P_x)$ is a point, a circle, a torus or a two-sphere (since it must be strongly homogeneous). We shall show, in the next section, that this is indeed the case. We shall then show that $G(x)$, one point x excepted, is a two-sphere.

6. Our analysis of $G(x)$ will depend upon the a priori possible structure of P_x. It is important to bear in mind that $G(x) = K_1(P_x)$ and that $G(x) - A \cdot G(x)$ is the product space of $P_x - A \cdot P_x$ and K_1 (see 4). Where details are intuitively clear and technically elementary, they will be omitted.

6.1 Suppose that P_x contains an arc qr having q and r only on A. In this, the typical case, $K_1(qr)$ is a two-sphere and must coincide with $G(x)$, by 4.1′. Furthermore, $P_x = qr$.

6.2 Suppose that P_x contains a simple closed curve C which has at least two points on A. Since C cannot be a subset of A we are led to the existence of an arc qr, as above, and then to a contradiction.

6.3 Suppose that P_x contains a simple closed curve C having no point on A. In this case (which will later be shown to be impossible) $K_1(C)$ is a torus and must coincide with $G(x)$. Furthermore, here, $P_x = C$.

6.4 Suppose that P_x contains a simple closed curve C which has precisely one point in common with A. In this case $K_1(C)$ is a pinch-torus which, separating space, must coincide with $G(x)$. This contradicts the homogeneity of $G(x)$ and is impossible.

6.5 Suppose that P_x is a tree one of whose endpoints is *not* in A. In this case it is clear that $K_1(P_x)$ is a set at least one of whose points is not an interior point of a two-cell. This contradicts section 5.2 and is impossible.

---

[10] We are not asserting, although it is the case, that these cells are open in $G(x)$.

6.6 Suppose that P_x is a tree not, degenerately, a single point. The tree must have at least two endpoints and, by the preceding discussion, both of these must lie in A. On the other hand, P_x cannot be a subset of A. Then we are led either to contradiction or to the situation of section 6.1.

We may summarize this section in the following:

6′) P_x is a point *or* it is an arc with endpoints (and these only) on A *or* it is a simple closed curve with no point on A.

6″) $G(x)$ is a point or circle *or* it is a two-sphere *or* it is a torus.

7. It is convenient to introduce at this point the continuous decomposition space R^* of E in which each orbit $G(x)$ is regarded as a "point" (4). This is a locally compact Peano space. We shall see that R^* must be a *ray* its single endpoint corresponding to the unique fixed point (in E) of G. It is clear that P may be regarded as a continuous decomposition space of E under the group K_1. But, further, the space R^* is a continuous decomposition space of P in which each set P_x is regarded as a point, namely the point $G(x) = K_1(P_x)$ of R^*. The last is an immediate consequence of the preceding statements.

7.1 It follows from this, but is also directly seen, that the points z of P for which P_z is a point form a closed set and that in consequence the points y for which P_y is a simple closed curve or an arc form an open and, as we know from sections 3 and 5.1′, non-vacuous subset of P. Now the "points" y^* of R^* corresponding to these points y must be cut-points of R^* of order two precisely. For, the surfaces which correspond to them in E are invariant under G and separate E into precisely two domains. The "points" y^* comprise all of the cut-points of R^*. For, the other "points" of R^* correspond to points z of P for which P_z is a point and $G(z)$ is an invariant circle or point; in either case a non-separating subset of E.

7.2 It follows at once from the Cyclic-Element theory (see e. g. 1 and 3) that R^*, being a Peano-space the set of whose cut-points is *open*, must be a tree. It must, furthermore, be a very simple tree since its cut-points are all of order two: in fact it can be an open curve (homeomorph of a line), a ray (homeomorph of a closed half-line), or an arc, and nothing else. On the other hand, each cutpoint of R^* must separate it into two components of which one is compact and the other is not. *For* that is precisely what the corresponding sphere or torus does in E. It is clear, then, that R^* is a *ray*.

7.3 Now the endpoint, let us call it f^*, of the ray R^* must correspond in E

to a point or to a circle, but in either event to a point f of P. This point f is the only point of P which satisfies the "equation" $P_x = x$. Now let a denote a point of $A - f$. Then $P_a \neq a$, and P_a must be an arc ab where b, also, is in A, by 6′). We shall prove, quickly, that f is a point of A between a and b. Let m_a denote a variable point of the arc ab of A and let it "move" continuously from a to b. In each position there is associated with it in a continuous way a point m_b which is the other endpoint of P_{m_a}. As m_a moves towards b the point m_b must advance towards a. This is immediately clear on a consideration of the half-plane P which is separated by each of these mutually exclusive arcs. It follows that in some position, $m_a = m_b$ and at this moment it must be the point f.

7. 4 Let R denote a definite one of the rays marked off on A by the point f. From the fact that R is closed and unbounded in E it is clear that the image of R in R^* cannot belong to any compact subset of R^*. This image, on the other hand, must contain f^* and must be connected. It follows that the image of R covers all of R^*. This implies, in particular, that all orbits $G(x)$ with the exception of $G(f) = f$ are two-spheres. Rather more than this is implied, however, for it is clear that R and R^* are, abstractly, entirely equivalent. In other words, R may be regarded as the decomposition space of E under the group G where the mapping is the one which carries each point x of E into the unique intersection of $G(x)$ with R. This will ultimately be seen to imply that the set of orbits is topologically equivalent to the family of spheres with center at the origin.

8. Now that orbits under G are seen to be two-spheres, as we wished them, we shall consider how G acts upon them individually. Throughout this section S will denote a particular one of these two-spheres and S is a $G(p)$ for some point p of R. Let T_p denote the subgroup of our original G' consisting of all elements for which the point p is a fixed point. This group contains K_1. We shall show that it coincides with K_1.

8. 1 By section 4. 1, $G'(p) = G(p)$. Then, if x denotes a point of S the orbit $T_p(x)$ is a subset of $G'(x) = S$ and contains $K_1(x)$. This last set is a simple closed curve, if we suppose that x is not the point p or that unique other point, call it $\bar{p}$, of $R \cdot S$. Then it follows, precisely as in section 4. 1, that $T_p(x) = K_1(x)$. From this, by the argument of section 3. 2, we conclude that T_p is identical with K_1. Therefore the set of points of S fixed under an arbitrary element of T_p, the identity excepted, consists of the points p and $\bar{p}$. For another point x, the orbit $T_p(x)$ is a simple closed curve.

8. 2 For the remainder of this section all symbols, x, y, z, etc., denoting

points will refer to points of S and the group G' will be regarded as a $Tg(S_2)$ where S_2 denotes the abstract two-dimensional sphere.

Let T_z denote the subgroup of G' leaving the point z fixed. Let y and g be such that $g(z)=y$. It is quite easy to see that $T_y = gT_zg^{-1}$, the transform of T_z by g. Then, if h is such that $h(p)=z$, $T_z = hT_ph^{-1}$ and T_z is a circle group. Let $\bar{z}$ denote $h(\bar{p})$, $\bar{p}$ as above. Clearly, $T_z(\bar{z})=\bar{z}$. Furthermore, from the fact that p and $\bar{p}$ only are fixed points of T_p it follows that z and $\bar{z}$ are the only fixed points of T_z. From the fact that $T_p = K_1 = T_{\bar{p}}$ (cf. section 8.1, replacing p by $\bar{p}$) it follows, at once, that $T_z = T_{\bar{z}}$. The point $\bar{z}$ will be called the *conjugate* of the point z. It is clear that the relation of conjugate is *unique* and *reciprocal*. Moreover, it is *invariant* under G'. For if T_x leaves x and $\bar{x}$ invariant, then gT_xg^{-1} leaves gx and $g\bar{x}$ invariant.

8.3 *The transformation $c(x)=\bar{x}$ is a homeomorphism of period two of S into itself.* Because of the remarks above, it suffices to prove the continuity of $c(x)$. Let $x_n \rightarrow x$ and let g_n be a sequence of elements of G' such that $g_n(x_n) = x$ *and* $g_n \rightarrow g_0$, the identity. Then, $g_n(\bar{x}_n)=\bar{x}$, by the last remark of the preceding paragraph, and $g_n^{-1}(\bar{x}) = \bar{x}_n$ must converge to $\bar{x}$.

8.4 For a fixed T_x all points, x and $\bar{x}$ excepted, have simple closed curves as orbits. The space of all these orbits, including $x = T_x(x)$ and $\bar{x} = T_x(\bar{x})$ is a closed interval and we may assign to each orbit $T_x(y)$ a real number t, $0 \leqq t \leqq 1$. This assignment is a continuous one, in a sense which is readily understood. We can now assign to each point y of S the same real number, let us call it $\alpha(y)$, which is associated with its orbit $T_x(y)$. We may "normalize" so that $\alpha(x)=0$ and $\alpha(\bar{x})=1$. From the continuity of this function it follows that if y is near x, $\alpha(y)$ is nearly 0 while $\alpha(\bar{y})$, in view of section 8.3, is nearly 1. Hence the function $\alpha(z)-\alpha(\bar{z})$ is negative near x and positive for points z near to $\bar{x}$. It follows that for at least one point y, $\alpha(y)=\alpha(\bar{y})$ and, for this y, $T_x(y)=T_x(\bar{y})$.

We have just seen that for every point x there is at least one orbit under the corresponding T_x which contains at least one pair of conjugate points. We shall show that this orbit is uniquely determined. For the present any such orbit will be called an *equator* with the point x as *pole* and denoted by Q_x.

8.5 *An equator, Q_x, contains the conjugate of each of its points.* For, let z belong to $Q_x = T_x(y)$ for some y. There is a g in T_x such that $g(y)=z$. But then $g(\bar{y})=\bar{z}$ and this point, too, belongs to Q_x. But, furthermore, the set Q_x divides S into two domains. The transformation $c(x)=\bar{x}$ having no fixed points and leaving Q_x invariant, as we have just seen, must interchange these domains. Accordingly, the only points w for which it can be

true that $T_x(w) = T_x(\bar{w})$ are points of Q_x. Therefore the equator Q_x is indeed unique and it is completely characterized by the fact that it is a self-conjugate orbit of T_x. Equivalently, the *equator may be characterized as a self-conjugate simple closed curve invariant under* T_x.

8.6 *The equator is an invariant of G' in this sense: if x, w, and g are such that $g(x) = w$ then $gQ_x = Q_w$.* For, first,

$$T_w(gQ_x) = gT_xg^{-1}(gQ_x) = gT_x(Q_x) = gQ_x$$

so that gQ_x is certainly invariant under T_w. Secondly, if z is a point of gQ_x $z = g(y)$ for some y of Q_x. Hence $\bar{z} = g(\bar{y})$ is also in gQ_x and, it follows, gQ_x is self-conjugate. Finally, it is clear that gQ_x is a simple closed curve. Therefore, by the preceding paragraph, gQ_x must be Q_w. It is a corollary to this that an equator is set-wise invariant under any element which interchanges it's "poles." For, if e is an element such that $e(x) = \bar{x}$, then

$$e(Q_x) = Q_{e(x)} = Q_{\bar{x}} = Q_x.$$

8.7 Now for any x there is at least one e such that $e(x) = \bar{x}$. For any such element, $e(\bar{x}) = x$ and $e^2(x) = x$. This element being in a connected group G' cannot reverse orientation on S but must interchange the domains determined on S by the invariant Q_x. Consequently it must reverse orientation on Q_x and it must have a pair of fixed points. These must, of course, be mutually conjugate. Let y denote one of them. Then $e(y) = y$, so that e is in T_y. Hence, since $e(x) = \bar{x}$, it follows that x is in Q_y for at least one y of Q_x. But let z denote any other point of Q_x. Then $z = h(y)$ for some h in T_x and $Q_z = hQ_y \supset h(x) = x$. *Then we have shown that for every point z on Q_x the "pole" x lies on Q_z.*

8.8 Here we shall show that *if an element of T_x interchanges one pair of conjugate points of Q_x it interchanges every pair on Q_x and is of order two.* Let g of T_x and y of Q_x be such that $g(y) = \bar{y}$. Then g^2 is in T_x and leaves y fixed. But in this case g^2 must be the identity of G' by 1′. Now let the integer n and the element h of T_x be such that $h^{2n} = g$. Then

$$gh(y) = h^{2n+1}(y) = hg(y) = h(\bar{y}) = \overline{h(y)}.$$

This means that g does interchange those pairs of conjugates which are of the form $h(y)$ and $h(\bar{y})$ for $h^{2n} = g$, $n = 0, 1, 2, \cdots$. But such points are everywhere dense on Q_x and our opening assertion follows from the continuity of the conjugate.

8.9 *We shall see now that if y and z are distinct non-conjugate points of Q_x then $Q_y \cdot Q_z = x + \bar{x}$.* From the last remark of section 8.7, $Q_y \cdot Q_z$ must contain x and $\bar{x}$. Suppose, then, that w is some point of this intersection. Let g denote an element of T_y for which $g(x) = \bar{x}$, and let h denote an element of T_z for which $h(x) = \bar{x}$. Since $hg(x) = x$ it is clear that hg is in T_x. On the other hand, $h(w) = \bar{w}$ and $g(w) = \bar{w}$ by section 8.8. Therefore, $hg(w) = w$. But if w is distinct from x and $\bar{x}$ this is possible if and only if hg is the identity. Now in this case, $y = hg(y) = h(y)$. Since h is an element of T_z and y is distinct from z and from $\bar{z}$, $y = h(y)$ implies that h is the identity. This contradiction to our choice of h concludes the argument.

8.10 *Let C be any simple closed curve orbit of T_x and let y be any point on Q_x. Then $C \cdot Q_y$ consists of precisely two points.* For, since C separates x and $\bar{x}$ while Q_y contains them, it follows that $C \cdot Q_y$ must contain at least two points a and b. There is some g in T_x for which $g(a) = b$. The point b must lie in gQ_y, since Q_y contains a. Then Q_y and $Q_{g(y)} = gQ_y$ have in common the point b which is certainly neither x nor $\bar{x}$. But this can be reconciled with section 8.9 only if $g(y) = y$ or $g(y) = \bar{y}$. The first alternative implies that g is the identity and is ruled out. The second alternative means, by section 8.8, that g^2 is the identity. Now since g is an element of the circle group T_x and is of order two, it is *uniquely* determined. Then it is clear that the intersection $C \cdot Q_y$ cannot contain more than two points and it must consist of the points a and $g(a)$ where g is the unique element of T_x of order two.

9. We are now in a position to set up on S and then to extend to E a coördinate system suitable for our purpose. This "spherical" coördinate system will be furnished by two appropriate subgroups which we shall call K and H, of G'. One of these we shall take as the group K_1, section 8.1, and now designate it by K. It would be possible to take the other as K_2 but it is more convenient at this point to forget K_2. We shall continue to use the notations of the preceding section.

Let q denote a definite point on the equator $Q_p = T_p(q)$ and let H denote the group T_q. The group $K = K_1$ is T_p; both H and K are circle groups. Let ϕ be a symbol for a group parameter specifying elements of the group H and, similarly, let θ be a symbol for elements of K. These symbols may be regarded as real numbers ranging from zero to one, and each group may be regarded as the additive group of the corresponding numbers, modulo one. We shall restrict our ϕ to the range: $0 \leqq \phi \leqq 1/2$ and denote the *set of these* ϕ's by H^*. Now $H^*(p)$ is an arc beginning at the point p and ending at the point $\bar{p}$ conjugate to p (section 8.2). The arc $H^*(p)$ is, of course, one of the two arcs $p\bar{p}$ on the equator Q_q.

We determine two coördinates θ_w and ϕ_w of an arbitrary point w of S as follows: the point w belongs to the unique orbit $K(w)$ and this meets the set $H^*(p)$ in precisely one point, by section 8.10. If we suppose that w is not p or $\bar{p}$, there is a unique parameter θ_w such that $\theta_w(w) = K \cdot Q_q$, $0 \leqq \theta_w < 1$. It is hoped that no confusion will arise here from our slightly ambiguous use of the parameter *as an element* of the group: we shall do this again with ϕ. If w is p or $\bar{p}$ then θ_w will be *chosen* as the parameter 0, corresponding to the identity of K. Now there is a unique parameter ϕ which we shall designate by ϕ_w, $0 \leqq \phi_w \leqq 1/2$, such that $\phi_w(p) = \theta_w(w)$: this parameter designates the unique element of H which carries p to the point $\theta_w(w)$ of Q_q.

To sum up, we have associated with each point w of S a unique pair of coördinates θ_w and ϕ_w such that, when we interpret these coördinates as the elements of G' with which they are associated, we have the following relation: $\theta_w^{-1}\phi_w(p) = w$.

We can now extend this coördinate system to the entire space E by introducing a third coördinate ρ: $0 \leqq \rho < \infty$ designating in a one-one continuous way the points of the axis R. The coördinate $\rho = 0$ corresponds to the unique fixed point f (section 7.4) under G'. This last coördinate does not, of course, correspond to any group-element of G'. Suppose, now, that w denotes an arbitrary point of E. If w is the point f, $\rho = 0$ and θ_w and ϕ_w will be defined to be 0. If w is not f the orbit $G'(w)$ is a two-sphere which has a unique point, call it p_w, on R. With this point there is associated a value of the coördinate ρ, call it ρ_w. Finally we determine the coördinates θ_w and ϕ_w by the observation that the sphere $G'(w) = G'(p_w)$ is completely and (the points p_w and $\bar{p}_w$ excepted) uniquely covered by the set of points $K\{H^*(p_w)\}$. That this is indeed the case is immediately clear if we regard $G'(p_w)$ as the sphere S of section 8 and following.

10. In this section we shall bring to a close the proof of our theorem. Let g denote an arbitrary element of G'. Let p be a point of R distinct from f and write $g(p) = w$. Now, $g(p) = \theta_w^{-1}\phi_w(p)$ and consequently, section 8.1, $g^{-1}\theta_w^{-1}\phi_w$ is an element of K. But then g^{-1}, hence g, is an element of that subgroup of G' which is generated by the two groups H and K. Therefore, G' itself must be the group generated by H and K. It is clear from this that, *in the coördinate system of* section 9, G' is the group of all "rigid rotations" of E_3 about the "origin" f.

SMITH COLLEGE, AND
QUEENS COLLEGE, NEW YORK.

BIBLIOGRAPHY.

1. Kuratowski and Whyburn, "Sur les éléments cycliques et leurs applications," *Fundamenta Mathematicae*, vol. 16, pp. 305-331.

2. Mayer and Thomas, "Foundations of the theory of Lie groups," *Annals of Mathematics*, vol. 36, pp. 770-822.

3. Menger, *Kurventheorie*, Berlin, 1932.

4. Montgomery and Zippin, "Periodic one-parameter groups in three-space," *Transactions of the American Mathematical Society*, vol. 40, pp. 24-36.

5. Montgomery and Zippin, "Compact abelian transformation groups," *Duke Mathematical Journal*, June, 1938.

6. Pontrjagin, "The theory of topological commutative groups," *Annals of Mathematics*, vol. 35, pp. 361-388.

7. Pontrjagin, "Théorie des groupes continus," *Comptes Rendus*, vol. 198, pp. 238-240.

A REMARK ON NORMAL EXTENSIONS.*

By O. F. G. Schilling.[1]

Let r be an arbitrary algebraic Riemann surface and G an arbitrary finite group. Then there exist covering surfaces R of r such that the group of relative automorphisms of R over r is isomorphic with G. The existence of such coverings R/r follows from Riemann's existence theorem.[2] In this note we want to present an existence theorem for arbitrary abstract fields of algebraic functions of one variable. The nature of our method, however, requires a restriction on the structure of the given finite group G.

Let G be a finite group which has the following properties. There exists a series of factor groups G_i of G such that

$$G_{i+1} = G_i/Z_i$$

where Z_i is a cyclic subgroup of G_i. We shall term G a pseudo-abelian group. Now let $g(u_{i+1})$, $g(v_{i+1})$, $g(u_{i+1}\, v_{i+1})$ denote a fixed set of representatives of G_i for the elements $u_{i+1}, v_{i+1}, u_{i+1}\, v_{i+1}$ of the factor group $G_i/Z_i = G_{i+1}$. Then

$$g(u_{i+1})\, g(v_{i+1}) = g(u_{i+1}\, v_{i+1}) z_i^{\alpha(u_{i+1}, v_{i+1})}$$

where z_i is a generator of the cyclic group Z_i of order m_i. The exponents $\alpha(u_{i+1}, v_{i+1})$ are integers which are uniquely determined by the selected representatives $g(u_{i+1})$, $g(v_{i+1})$, $g(u_{i+1}\, v_{i+1})$ and the generator z_i.

Let $k = \Omega(x, y)$ be a field of algebraic functions of one variable whose field of constants Ω is algebraically closed. Let χ be the characteristic of Ω. Suppose now that we already constructed a normal extension K of k whose Galois group is isomorphic with a given pseudo-abelian group G. The assumptions on G imply that K contains a series of normal subfields K_i/k corresponding to the factor groups G_i. Moreover, K_i/K_{i+1} is a cyclic extension and its Galois group is equal to Z_i. Thus, every field K_{i+1} with the Galois group $G_{i+1} = G_i/Z_i$ is imbedded in a normal field K_i whose Galois group is equal to G_i.

Theorem 1. *Let $k = \Omega(x, y)$ be a function field of characteristic χ and*

* Received October 24, 1938.

[1] Johnston Scholar at the Johns Hopkins University for the year 1938-1939.

[2] A. Weil, "Généralisation des fonctions abéliennes, *Journal de Liouville,* vol. 17 (1937), ser. 9.

G a pseudo-abelian group whose order is relatively prime to χ. There exist infinitely many normal extensions K over k whose Galois groups are isomorphic with G.

Proof. We proceed by induction. Let $G_r = Z_r \neq 1$ be the last factor group in a composition of the given pseudo-abelian group G. Since the order of G is relatively prime to the characteristic χ of k we can find an infinitude of radical extensions

$$K_r = k(a^{1/m_r})$$

of degree m_r over k. The Galois group of any one of these fields K_r/k is isomorphic with $G_r = Z_r$. Let us take a particular one of these fields K_r and suppose that we already constructed a normal extension K_{i+1}/k whose Galois group is isomorphic with $G_{i+1} = G_i/Z_i$. We want to prove the existence of a normal extension K_i which contains K_{i+1} and whose Galois group over k is isomorphic with G_i. We express this problem in terms of the theory of linear algebras. The relations

$$g(u_{i+1})\, g(v_{i+1}) = g(u_{i+1}\, v_{i+1}) z_i^{a(u_{i+1}, v_{i+1})}$$

which define the group G_i in terms of Z_i and G_{i+1} give rise to a factor set of K_{i+1} relative to k. We define the factor set by

$$a(u_{i+1}, v_{i+1}) = \zeta_i^{a(u_{i+1}, v_{i+1})}$$

where ζ_i denotes a fixed primitive m_i-th root of unity. A theorem of R. Brauer yields that K_{i+1} can be imbedded in a normal field K_i in the prescribed fashion if the algebra

$$A_i = (K_{i+1}/k,\ G_{i+1},\ \zeta_i^{a(u_{i+1}, v_{i+1})})$$

is a full matric algebra over k.[3] Since the field of constants Ω is algebraically closed, it follows that

$$A_i \sim k.^{4}$$

Consequently the existence of a normal field K_i is established.

COROLLARY. *In particular, there exist infinitely many fields K over k whose Galois groups are groups of order l^m where $(l, \chi) = 1$.*[5]

---

[3] R. Brauer, "Über die Konstruktion der Schiefkörper, die von endlichen Rang in bezug auf ein gegebenes Zentrum sind," *Crelle*, vol. 168 (1932).

[4] C. C. Tsen, "Divisionsalgebren über Funktionenkörpern," *Göttinger Nachrichten* (1934).

[5] A. Speiser, "Theorie der Gruppen endlicher Ordnung" (2nd edition), Theorem 78, p. 69.

We now want to prove a refinement to Theorem 1. Let G be a group of order l^m. Then the totality of elements $g_1 g_2 g_1^{-1} g_2^{l-1}$, where $g_1, g_2 \in G$, generates an invariant subgroup G^* of G. The factor group G/G^* is an abelian group of type $(l, l, \cdots, l)$; let l^n be its order. Then n is equal to the least number of generators of the group G.

Let K be a normal extension of k whose Galois group is equal to G. Then there belongs to $G^* \subset G$ a subfield K^* of K which is the join of all cyclic subfields Z_l of degree l over k. We can build up K over k in the following fashion:

$$k \subset K^* \subset K_1 \subset \cdots \subset K_j \subset K_{j+1} \subset \cdots \subset K$$

where the fields K_j are normal over k and K_{j+1} are cyclic extensions of degree l over K_j.

LEMMA 1. *Let L be a normal extension of degree l^m over $k = \Omega(x, y)$ and suppose that some prime divisor of k is completely ramified in L. Then L is a cyclic extension of k provided that $(l, \chi) = 1$.*

Proof. Let $\mathfrak{p}$ be the completely ramified prime divisor of k. Then $\mathfrak{p} = \mathfrak{P}^{l^m}$ in L. The ramification theory implies that the ramification group of $\mathfrak{P}$ coincides with the full Galois group G of L over k. Consequently the Galois group of the complete field $L_{\mathfrak{P}}$ over $k_{\mathfrak{p}}$ is equal to G. But $L_{\mathfrak{P}} = k_{\mathfrak{p}}(\pi^{1/l^m})$ where $v_{\mathfrak{p}}(\pi) = 1, \pi \in k_{\mathfrak{p}}$. Namely $(l, \chi) = 1$ according to assumption. Hence G is cyclic.

LEMMA 2. *If $L = k(a^{1/l})$ and $a = \Pi \mathfrak{p}_i^{a_i}$ then the discriminant of L over k contains exactly those prime divisors $\mathfrak{p}_i$ for which $v_{\mathfrak{p}_i}(a) \not\equiv 0(l)$.*

Proof. The discriminant of L over k is a divisor of the discriminant of the defining equation $y^l - a = 0$. The latter is equal to $(-1)^{l(l-1)/2} l^l (-a)^{l-1}$. Since $(l, \chi) = 1$ it follows that at most the prime factors $\mathfrak{p}_i$ of a can be ramified in L. Let $\mathfrak{p}$ be such a prime divisor and $k_{\mathfrak{p}}$ the associated $\mathfrak{p}$-adically closed field of k. Then

$$L \times k_{\mathfrak{p}} = k_{\mathfrak{p}} \dotplus \cdots \dotplus k_{\mathfrak{p}} \quad \text{if } v_{\mathfrak{p}}(a) \equiv 0(l)$$

and

$$L \times k_{\mathfrak{p}} = L_{\mathfrak{P}} \qquad \text{if } v_{\mathfrak{p}}(a) \not\equiv 0(l).$$

Both assertions immediatey follow from the ramification theory. Thus the Lemma is proved.

Let $\tau(l)$ be the number of divisor classes of degree 0 in k whose exponents are equal to l. The number $\tau(l)$ is finite as a simple algebraico-geometric

argument yields.[6] The maximal unramified abelian extension of exponent l over k has then the degree $l^{\tau(l)}$.

Now let $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ be a finite number of prime divisors of k. We want to determine the degree of the maximal abelian extension K_l over k,

whose exponent is equal to l, and
which is ramified at most at $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$.

LEMMA 3. *The maximal abelian extension K_l which has exponent l and is ramified at most at $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ has degree $l^{\tau(l)+r-1}$ over k.*

Proof. The arithmetical theory of radical extensions yields that K is the join of cyclic extensions $Z_i = k(a_i^{1/l})$. The numbers $a_i \in k$ are described by the property that

$$v_{\mathfrak{p}_j}(a_i) \not\equiv 0(l)$$

for at most $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$. (See Lemma 2.) Hence we have to determine the index $[a : b^l]$ where

$$v_{\mathfrak{p}_j}(a) \not\equiv 0(l), \qquad (j = 1, \cdots, r),$$

a are not l-th powers, and
b runs over all elements different from 0 in k.

Let $\mathfrak{A}$ be the group of all divisors in k and c the group of all numbers in k for which

$$v_{\mathfrak{p}}(c) \equiv 0(l) \quad \text{for all } \mathfrak{p} \text{ of } k.$$

Moreover let $\mathfrak{A}' = \{\mathfrak{a}'\}$ the group of all divisors in k for which at most

$$v_{\mathfrak{p}_j}(\mathfrak{a}') \equiv 0(l) \quad \text{for } \mathfrak{p}_j \in \{\mathfrak{p}_1, \cdots, \mathfrak{p}_r\}.$$

Then

$$[a : b^l] = [\mathfrak{A}' : \mathfrak{A}^l][c : b^l][\mathfrak{A} : \mathfrak{A}^l(b)]^{-1}[\mathfrak{A} : \mathfrak{A}'(b)]$$

according to the principles of reduction which are familiar in class field theory.[7] Therefore

$$[a : b^l] = l^{r-1} l^{\tau(l)}[c : b^l][\mathfrak{A} : \mathfrak{A}^l(b)]^{-1}$$

according to the structure of the group $\mathfrak{A}$. Moreover,

$$[c : b^l] = [\mathfrak{A} : \mathfrak{A}^l(b)]$$

---

[6] O. F. G. Schilling, "Foundations of an abstract theory of abelian functions," *American Journal of Mathematics*, vol. 61 (1939).

[7] H. Hasse, "Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper" (Leipzig, 1930), Part Ia, §§ 15, 16.

for the same reason as before. Consequently,

$$[c : b^l] = l^{r+\tau(l)-1}$$

and the a_i can be chosen as a complete set of representatives of the factor group c/b^l.

THEOREM 2. *Let G be a group of order l^m whose least number of generators is equal to n. There always exist normal extensions K over k whose Galois group is isomorphic with G provided that $n \leqq r + \tau(l) - 1$. The necessary r branch points $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ can be prescribed arbitrarily.*

Proof. Let r branch points $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ be chosen such that $n \leqq r + \tau(l) - 1$. Suppose moreover that $r = n - \tau(l) + 1$. We want to construct a normal extension K over k which is at most ramified at $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$. According to Lemma 3 we can construct an abelian extension K_l/k of exponent l which is at most ramified at $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$. We find that $[K_l : k] = l^{r+\tau(l)-1}$. In order to construct the field K we put $K^* = K_l$. Then we get the right number of generators. Theorem 1 implies the existence of fields K_j such that

$$k \subset K^* \subset K_1 \subset \cdots \subset K_j \subset K_{j+1} \subset \cdots \subset K$$

and such that K has the right Galois group over k. In order to prove that at most $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ are ramified in K too, we again use induction. Suppose we already constructed a normal extension K_j/k. We want to show that K_{j+1} can be chosen such as to have no other branch points than the given $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$. Suppose now that $\mathfrak{p}$ is a branch point of some constructed imbedding field K_{j+1} which does not occur amongst the given set $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$. Lemma 3 yields the existence of a cyclic field Z of degree l over k which is ramified at $\mathfrak{p}$ and at least at one of the $\mathfrak{p}_j$. Namely, we only have to consider the maximal field K'_l belonging to $\mathfrak{p}, \mathfrak{p}_1, \cdots, \mathfrak{p}_r$. The field Z is surely not contained in K^*. Consequently the join $K_{j+1}Z$ is an abelian extension of degree l^2 over K_j. The Galois group of $K_{j+1}Z$ over K_j has type (l, l) and the Galois group of $K_{j+1}Z$ over k is equal to the direct product of G_{j+1} and a cyclic group of order l. Since $\mathfrak{p}$ does not divide the discriminant of K_j over k we have

$$\mathfrak{p} = \mathfrak{P}^*_1 \cdots \mathfrak{P}^*_{l^t}$$

in K_j where $[K_j : k] = l^t$. A prime divisor P^*_i of K_j decomposes in $K_{j+1}Z$ as follows:

$$\mathfrak{P}^*_i = (\mathfrak{P}_{i,1} \cdots \mathfrak{P}_{i,l})^l.$$

Namely $\mathfrak{P}^*_i = \mathfrak{P}_i^{l^2}$ would imply according to Lemma 2 that $K_{j+1}Z$ is a cyclic extension of degree l^2 over K_j.

The theory of the decomposition groups of $\mathfrak{P}^*_i$ relative to $K_{j+1}Z$ and K_j implies that $\mathfrak{P}^*_i$ is totally decomposed in one of the fields $L(\neq K_{j+1}Z, K_j)$ lying between $K_{j+1}Z$ and K_j. The structure of the Galois group of $K_{j+1}Z$ over k yields that the Galois group of L over k is isomorphic with G_{j+1}.[8] Moreover $\mathfrak{p}$ is not ramified in L for $\mathfrak{P}^*_i$ is not ramified in L, and L is a normal field. Repeating this process a finite number of times we finally arrive at a field $L = K'_{j+1}$ which is at most ramified at $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$. Since $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ are ramified in K^* the field K which we ultimately obtain, has the desired properties.

It is now obvious that one can also construct normal fields K which are exactly ramified at $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ provided that $n \leqq r + \tau(l) - 1$.

Remark 1. If Ω is an algebraically closed field of characteristic 0 then $\tau(l) = 2p$ where p denotes the genus of k. Thus Theorem 3 yields $n \leqq 2p + r - 1$ which is the formula obtained in the classical theory.

Remark 2. It is not difficult to compute the exact number of fields K which are exactly ramified at given prime divisors $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$. The necessary calculations are similar to the ones given by Witt.[9]

Remark 3. If $k = \Omega(x)$ then we must have at least $n + 1$ branch points.

Let Ω be a finite Galois field of $q = \chi^\nu$ elements. Suppose moreover that Ω contains the l-th roots of unity, i. e., $q \equiv 1 \pmod{l}$. Under these assumptions it is possible to generalize Theorem 1. However, it is necessary to impose special conditions upon the prime divisors $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ of $k = \Omega(x, y)$ which shall be ramified in the normal extension K/k. Namely, in order to be able to apply the theory of algebras, it is necessary to insure by appropriate assumptions on the given prime divisors $\mathfrak{p}_i$ that the algebras

$$A_i = (K_i/k, \cdot G_i, \zeta_i^{a(u_{i+1}, v_{i+1})})$$

are similar to k. Since the norm theorem is valid for the cyclic extensions over k, it must be proven that

$$A_i \times k_{\mathfrak{p}_j} \sim k_{\mathfrak{p}_j}$$

for every $\mathfrak{p}_j$. The discriminant of A_i contains at most the ramified prime

[8] H. Reichardt, "Konstruktion von Zahlkörpern mit gegebener Galoisgruppe von Primzahlpotenzordnung," *Crelle*, vol. 177 (1937).

[9] E. Witt, "Konstruktion von galoisschen Körpern der Charakteristik p mit gegebener Gruppe der Ordnung p^f," *Crelle*, vol. 174 (1936).

divisors of K_i/k, the factor set giving no contribution.[10] In order to insure $A_i \times k_{\mathfrak{p}_j} \sim k_{\mathfrak{p}_j}$ it suffices to suppose that the residue class fields of the different $\mathfrak{p}_j$ contain sufficiently many l^ρ-th roots of unity. In other words, we must suppose that $q^{f(\mathfrak{p}_j)} \equiv 1(l^\rho)$, where $f(\mathfrak{p}_j)$ denotes the absolute degree of $\mathfrak{p}_j$. This condition can always be realized. Moreover, we shall require that the maximal abelian subfield K^* of the field K to be constructed be completely ramified. This means that $r = n$. Having satisfied these conditions, we are able to state the following existence theorem whose proof can be taken over verbatim from the theory of algebraic number fields.[11]

THEOREM 3. *Let G be a finite group of order l^m whose least number of generators is equal to n. Suppose that there are given n prime divisors $\mathfrak{p}_1, \cdots, \mathfrak{p}_n$ of $k = \Omega(x, y)$ whose absolute degrees $f(\mathfrak{p}_i)$ are sufficiently large so as to insure $q^{f(\mathfrak{p}_i)} \equiv 1(l^m)$. There exists a normal extension K over k whose Galois group is isomorphic with G and which is ramified exactly at the given n prime divisors.*

Remark. Let $l^{\tau(l)}$ be the degree of the maximal abelian unramified extension of exponent l over k which does not contain the cyclic extension of degree l over Ω. A careful analysis of the proof of Theorem 3 yields that the theorem on unramified normal extensions generalizes. Namely, if the number of generators of the given group G is less or equal than $\tau(l)$ then there exist normal unramified fixed K whose Galois group is isomorphic with G.

We remark that the existence of a normal field K with the Galois group G over k yields in general no information on the nature of the ramification. Namely, the theorem on the structure of algebras which we used is a sufficient (but not necessary) condition.

We next want to discuss the case of groups G whose orders l^m are powers of the characteristic χ. The general existence proof for such groups was given by E. Witt.

Let $k_\mathfrak{p}$ be the p-adic closure of $k = \Omega(x, y)$ with regard to the prime divisor k. Then every cyclic extension Z of $k_\mathfrak{p}$ of degree χ is given by the root of an equation

$$z^\chi - z = a_\mathfrak{p}.$$

[10] H. Reichardt, "Die Diskriminante einer normalen einfachen Algebra," *Crelle*, vol. 173 (1935).

[11] T. Tannaka, "Zyklische Zerfällungskörper der einfachen Ringe über dem algebraischen Funktionenkörper," *Sci. Rep. Tôhoku Univ.*, vol. XXIV (1935).

T. Tannaka, Über die Konstruktion der galoisschen Körper mit vorgegebener p-Gruppe. *Tôhoku Mathematical Journal*, vol. 43 (1937).

Let $a_{\mathfrak{p}} = \omega\pi^{-\lambda} + \cdots$ where $\omega \epsilon \Omega$, $v_{\mathfrak{p}}(\pi) = 1$ and $\lambda \geqq 0$. Since Ω is supposed to be algebraically closed we always get $\lambda > 0$, $\lambda \not\equiv 0(\chi)$. The discriminant of $Z = k_{\mathfrak{p}}(z)$ is equal to $\mathfrak{p}^{(\chi-1)(\lambda+1)}$. Consequently there exist infinitely many different cyclic extensions Z of degree χ over $k_{\mathfrak{p}}$. Therefore the index $[a : b^{\chi} - b]$ is infinite. Applying Witt's result we get the following theorem.

THEOREM 4. *Let G be a group of order χ^m and $k_{\mathfrak{p}}$ a complete discrete field of characteristic χ. There exist infinitely many normal extensions K over $k_{\mathfrak{p}}$ whose Galois groups are equal to G.*

Remark. Suppose that Ω is a finite Galois field. If we require that the discriminants of the extensions K over $k_{\mathfrak{p}}$ are bounded, we obtain restrictions on the number of generators n of G. Moreover, one finds that there exists only a finite number of such fields. The computations are simple and similar to the ones carried out by Witt. The actual computations show that the number of generators depends on the given bound and the number of elements in Ω.

In order to determine the totality of all cyclic extensions Z of degree χ over $k = \Omega(x, y)$ which are at most ramified at a finite number of prime divisors $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$, we have to consider the index $[a' : b^{\chi} - b]$. There a' denotes the additive group of all elements a' in k for which

$$v_{\mathfrak{p}_j}(a') \not\equiv 0(\chi) \quad \text{and} \quad v_{\mathfrak{p}_j}(a') < 0, \qquad (j = 1, \cdots, r).$$

Witt's results imply that the index $[a' : b^{\chi} - b]$ is infinite as soon as $r \geqq 1$.

THEOREM 5. *Let G be a group of order χ^m. There exist infinitely many normal extensions K of $k = \Omega(x, y)$ whose Galois groups are isomorphic with G and which are ramified at a single prime divisor $\mathfrak{p}$ of k.*

Proof. Let again G^* be the subgroup of G which is generated by all elements of the form $g_1 g_2 g_1^{-1} g_2^{\chi-1}$. Then G/G^* is an abelian group of order χ^n and type $(\chi, \chi, \cdots, \chi)$. Let $\mathfrak{p}$ be the given prime divisor. Then there exists at least one abelian extension K^* over k which is at most ramified at $\mathfrak{p}$. The field K we have to construct is determined by a chain of cyclic extensions K_{j+1}/K_j of degree

$$k \subset K^* \subset K_1 \subset \cdots \subset K_j \subset K_{j+1} \subset \cdots \subset K.$$

According to Witt there exist fields K_j such that K has the right Galois group over k. A relatively simple argument using Theorem 3 and the theory of divisor classes yields that the extensions K_j can be chosen such that no new

ramified prime divisors except the divisors of $\mathfrak{p}$ are added. The classes to be considered are $(\mathfrak{P}_j{}^t)$, $t > 0$, where $\mathfrak{P}_j \mid \mathfrak{p}$ in K_j.

Remark 1. Since k contains only a finite number of divisor classes of degree 0 and exponent χ the number of generators n of G is limited if we want to construct unramified extensions K/k.[12] One actually can exhibit examples of fields k of genus p for which $\tau(\chi) = 1$. It is never possible to construct unramified extensions K/k whose Galois groups have order χ^m.

Remark 2. Let $k = \Omega(x)$ be a rational function field and G a group of order χ^m. There exist always infinitely many normal fields K/k whose Galois groups are isomorphic with G and which are ramified at a single prescribed prime divisor $\mathfrak{p}$ of k. If we impose the condition that the discriminant of K/k is bounded and that Ω is finite, then we can again find relations between the number of generators and the bound. The number of different fields can be determined; it is finite.

Remark 3. If Ω is a finite Galois field of $q = \chi^\nu$ elements, then most of the preceding results can be carried over mutatis mutandis. It turns out that the number q and the genus p of the underlying field k have to be considered. But we do not insist on an explicit computation of the respective formulae.[13]

Finally we remark that the preceding theorems when compared with the theory of algebraic number fields, indicate the close relationship between the genus of our function fields and the class number of an algebraic number field.

THE JOHNS HOPKINS UNIVERSITY.

---

[12] H. Hasse and E. Witt, "Zyklische unverzweigte Erweiterungskörper vom Primzahlgrade p über einem algebraischen Funktionenkörper der Charakteristik p.," *Monatshefte für Mathematik und Physik*, vol. 43 (1936).

[13] E. Witt, "Der Existenzsatz für abelsche Funktionenkörper," *Crelle*, vol. 173 (1935).

BOREL SUMMABILITY AND LAMBERT SERIES.*

By Tomlinson Fort.

There are several theorems in the theory of convergent series of which the following [1] is an example.

If $a_n(z)$ and $b_n(z)$, $n = 0, 1, 2, \cdots$, are defined over a set P and if

$$\sum_{n=0}^{\infty} a_n(z) \tag{1}$$

converges uniformly over P and

$$\sum_{n=0}^{\infty} | b_{n+1}(z) - b_n(z) | \leqq G,$$

a constant, and $| b_0(z) | < b$, a constant; then

$$\sum_{n=0}^{\infty} a_n(z) b_n(z)$$

converges uniformly over P.

We shall refer to these theorems as *introductory theorems.* They are a ready consequence of the formula for summation by parts,

$$\sum_{n=k}^{n} u_n w_n = u_{n+1} \sum_{n=k}^{n} w_n - \sum_{n=k}^{n} (\Delta u_n) \sum_{n=k}^{n} w_n. \tag{2}$$

In the present paper theorems analogous to *introductory theorems* are proved for Borel Integral summability. The theorems which are proved are applied to Lambert series.

1. Borel's integral definition.[2]

We shall prove the following theorem:

Theorem I. *If $a_n(z)$ and $b_n(z)$, $n = 0, 1, 2, \cdots$, are defined at all points of a set P, $b_n(z) \to b(z)$ where $| b(z) | < B$ and if* (1) *is uniformly Borel integral summable over P; if, moreover, there exists a function $f(n, z) \neq 0$ when z is of P such that*

---

* Received September 12, 1938.

[1] The names Abel, Dedekind, Dirichlet and du Bois-Reymond have been attached to theorems of this type. See Knopp, *Theorie und Anwendung der unendlichen Reihen,* S. 316; Bromwich, *Theory of Infinite Series* (Second Edition), p. 246; Fort, *Infinite Series,* p. 106.

[2] *Leçons sur les séries divergentes,* ed. 1901, p. 98.

$$\sum_{n=0}^{\infty} |\Delta b_n(z)|\ |f(n,z)| \tag{3}$$

converges[3] *uniformly in* z *over* P *to a bounded function and such that*

$$\frac{1}{f(n,z)} e^{-\alpha} \sum_{n=0}^{n} a_n(z) \frac{\alpha^n}{n!} \tag{4}$$

is bounded in n *and* z *and such that*[4]

$$\int_0^{\infty} \frac{1}{f(n,z)} e^{-\alpha} \sum_{n=0}^{n} a_n(z) \frac{\alpha^n}{n!}\, d\alpha$$

converges uniformly in n *and* z; *then*

$$\sum_{n=0}^{\infty} a_n(z) b_n(z) \tag{5}$$

is uniformly Borel integral summable over P.

Denote the Borel integral formed for (1) by s. Then

$$s = \int_0^{\infty} e^{-\alpha} \sum_{n=0}^{\infty} b_n(z) a_n(z) \frac{\alpha^n}{n!}\, d\alpha. \tag{6}$$

Now in (2) let $k = 0$, $u_n = b_n$ and $w_n = a_n\alpha^n/n!$ and substitute in (6). We get

$$s = \int_0^{\infty} \lim_{n\to\infty} \Big[b_{n+1}(z) e^{-\alpha} \sum_{n=0}^{n} a_n(z) \frac{\alpha^n}{n!} - \sum_{n=0}^{n} (\Delta b_n(z)) e^{-\alpha} \sum_{n=0}^{n} a_n(z) \frac{\alpha^n}{n!} \Big]\, d\alpha. \tag{7}$$

Denote the integrand in (6) by $q(\alpha, z)$. We have

$$\left| \int_k^{k'} q(\alpha,z)\, d\alpha \right| < B \left| \int_k^{k'} e^{-\alpha} \sum_{n=0}^{n} a_n(z) \frac{\alpha^n}{n!}\, d\alpha \right| + \left| \int_k^{k'} \sum_{n=0}^{\infty} (\Delta b_n(z)) e^{-\alpha} \sum_{n=0}^{n} a_n(z) \frac{\alpha^n}{n!}\, d\alpha \right|.$$

Since (4) is bounded in n, α and z and since (3) converges uniformly in z over P, the infinite series under the second integral sign in the right-hand member converges uniformly in α by the Weierstrass[5] test, and can be integrated term by term. Consequently, we can write

---

[3] $\Delta b_n(z) = b_{n+1}(z) - b_n(z)$.

[4] Throughout this paper α is real.

[5] Fort, *Infinite Series*, Theorem 113.

$$\left| \int_k^{k'} q(\alpha, z)\, d\alpha \right| < B \left| \int_k^{k'} e^{-\alpha} \sum_{n=0}^{\infty} a_n(z) \frac{\alpha^n}{n!}\, d\alpha \right|$$
$$+ \left| \sum_{n=0}^{\infty} (\Delta b_n(z)) f(n, z) \int_k^{k'} \frac{1}{f(n, z)} e^{-\alpha} \sum_{n=0}^{n} a_n(z) \frac{\alpha^n}{n!}\, d\alpha \right|.$$

We have assumed that (1) is uniformly Borel integral summable over P, that is, that

$$\int_0^{\infty} e^{-\alpha} \sum_{n=0}^{\infty} a_n(z) \frac{\alpha^n}{n!}\, d\alpha$$

is a uniformly convergent integral. Under this assumption given any $\eta > 0$ choose k so large that when $k' > k > K$, both

$$\left| \int_k^{k'} e^{-\alpha} \sum_{n=0}^{\infty} a_n(z) \frac{\alpha^n}{n!}\, d\alpha \right| < \eta$$

and

$$\left| \int_k^{k'} \frac{1}{f(n, z)} e^{-\alpha} \sum_{n=0}^{n} a_n(z) \frac{\alpha^n}{n!}\, d\alpha \right| < \eta.$$

Then

$$\left| \int_k^{k'} q(\alpha, z)\, d\alpha \right| < \eta[B + \sum_{n=0}^{\infty} |\Delta b_n(z)|\ |f(n, z)|].$$

Since by hypothesis

$$\sum_{n=0}^{\infty} |\Delta b_n(z)|\ |f(n, z)|$$

is bounded

$$\left| \int_k^{k'} q(\alpha, z)\, d\alpha \right| < \eta(B + B') < \epsilon$$

if $\eta < \epsilon/(B + B')$.

Under the assumptions made we have thus established the uniform convergence of the integral (6) and so the uniform Borel integral summability of (5) and so Theorem I.

2. Absolute summability.[6] Theorem 1 admits of an immediate analogue for absolute Borel integral summability. Proof is omitted to avoid repetition.

THEOREM II. *If $a_n(z)$ and $b_n(z)$ are defined at all points of P, $b_n(z) \to b(z)$ where $|b(z)| < B$ and if (1) is absolutely and uniformly Borel integral summable over P, moreover, if there exists a function $f(n, z, \lambda) \neq 0$ when z is of P such that*

$$\sum_{n=0}^{\infty} |\Delta b_n(z)|\ |f(n, z, \lambda)|$$

[6] Borel, *loc. cit.*, p. 99.

converges uniformly over P to a bounded function for each λ and such that

$$\frac{1}{f(n,z,\lambda)}\, e^{-\alpha} \sum_{n=\lambda}^{n} a_n(z)\, \frac{\alpha^{n-\lambda}}{(n-\lambda)\,!}$$

is bounded in n, z and α for each λ and

$$\int_0^{\infty} \frac{1}{|f(n,z,\lambda)|}\, e^{-\alpha} \left| \sum_{n=\lambda}^{n} a_n(z)\, \frac{\alpha^{n-\lambda}}{(n-\lambda)\,!} \right| d\alpha$$

is uniformly convergent in n and z over P for each λ, then (5) is absolutely and uniformly Borel integral summable over P.

3. Lambert and Power Series. Consider the power series

(8) $$\sum_{n=0}^{\infty} d_n z^n.$$

Denote the interior of the Borel polygon of summability [7] of (8) by S.

We shall proceed to prove the following theorem.

THEOREM III. *Let T be the radius of convergence of (8). Let $T' < T$ and let $f(n,z,\lambda) = |z/T'|^n$ if $|z| > T$ and $f(n,z,\lambda) = 1$ if $|z| \leqq T$. Then series (8) fulfills all the conditions required of the a-series in Theorem II so long as z is restricted to a subregion R of S lying in the finite plane and with boundary always distant from the boundary of S by $\bar{\epsilon}$.*

Let

$$q(\alpha,z) = \sum_{n=0}^{\infty} d_n \frac{z^n \alpha^n}{n\,!}$$

and let

$$q_\lambda = \frac{\partial^\lambda q(\alpha,z)}{\partial \alpha^\lambda}.$$

Borel proved [8] that when z lies in R.

(10) $$|q_\lambda| < k(\lambda) e^{(1-\epsilon)\alpha}$$

where $k(\lambda)$ is independent of α and z. We proceed to examine

$$s_n(\alpha,z,\lambda) = \sum_{n=\lambda}^{n} d_n \frac{z^n \alpha^{n-\lambda}}{(n-\lambda)\,!}.$$

We write

(11) $$q_\lambda = s_n(\alpha,z,\lambda) + r_n(\alpha,z,\lambda).$$

---

[8] Borel, *loc. cit.*, p. 124.

[7] Borel, *loc. cit.*, p. 128.

Draw a simple closed curve, c, lying wholly within the polygon of summability, enclosing R and distant from the boundary of the polygon of summability by ϵ and from the boundary of R by δ. Let L be the length of c. Then [9]

$$(12) \qquad r_n(\alpha, z, \lambda) = z^n \int_c \frac{q_\lambda dt}{t^n(t-z)}.$$

From this and from (10)

$$(13) \qquad | r_n(\alpha, z, \lambda) | \leqq e^{(1-\epsilon)\alpha} k(\lambda) \frac{| z |^n}{\delta(T-\delta)^n} = e^{(1-\epsilon)\alpha} K(\lambda) \frac{| z |^n}{T'^n}.$$

Hence by (10), (11) and (13)

$$(14) \qquad | s_n(\alpha, z, \lambda) | \leqq \bar{K}(\lambda) \left(1 + \frac{| z |^n}{T'^n}\right) e^{(1-\epsilon)\alpha} \leqq 2\bar{K}(\lambda) \frac{| z |^n}{T'^n} e^{(1-\epsilon)\alpha}.$$

if $| z | > T$ and

$$(15) \qquad | s_n(\alpha, z, \lambda) | \leqq 2\bar{K}(\lambda) e^{(1-\epsilon)\alpha}$$

if $| z | \leqq T'$. The theorem follows from (10), (14) and (15).

We now consider the Lambert Series

$$(16) \qquad \sum_{n=1}^{\infty} d_n \frac{p^n z^n}{p^n - z^n}.$$

It can readily be proved by the *introductory theorems* that this converges at all points at which (8) converges for which $| z | \neq | p |$ and absolutely uniformly over any finite region lying within the circle of convergence of (8) and whose boundary remains a distance δ from the circle of convergence and the circle $| z | = | p |$.

We proceed to examine the conditions imposed on the b-series by Theorem II when $b_n = p^n/(p^n - z^n)$. We announce the following theorem.

THEOREM IV. *If T is the radius of convergence of* (8) *and $T' < T$ and $f(n, z, \lambda) = | z |^n/T'^n$ if $| z | > T$ and $f(n, z, \lambda) = 1$ if $| z | \leqq T$, then $b_n = p^n/(p^n - z^n)$ fulfills all the conditions imposed on the b-series by Theorem* II *if z is at least distant from the circle $| z | = | p |$ by ϵ and either $| z | \leqq T'$ or $| z | < p$ and $| z^2 | < | pT' |$ simultaneously.*

This theorem is readily proved by means of the test-radio test of the elementary theory of series. We, consequently, are able to state the following theorem.

[9] See for example, Osgood, *Funktionentheorie* (zweite Auflage), S. 316.

THEOREM V. *Series* (16) *is absolutely summable by the Borel integral method at all points at which series* (8) *is absolutely Borel integral summable provided* $|z| \neq |p|$ *and either* $|z| < T$, *the radius of convergence of* (8), *or* $|z| < |p|$ *and* $|z^2| < |pT|$ *simultaneously. It is uniformly summable over any bounded region within the interior of the polygon of summability of* (8) *such that* $|z| \leqq |p| - \epsilon$ *and* $|z^2| < |p|(T - \epsilon)$ *simultaneously at all points of the region.*

It is to be remarked that if $|z| < T$ and $|z| \neq |p|$, (16) is absolutely convergent and hence absolutely summable.

It is further to be remarked that on account of the symmetry of (16) any theorem relative to z, p is also true relative to p, z.

LEHIGH UNIVERSITY.

SINGULAR MAPS OF DIFFERENTIABLE MANIFOLDS OF N DIMENSIONS INTO N DIMENSIONAL EUCLIDEAN SPACE.*

By T. Y. Thomas.

1. Let M be a separable and compact differentiable manifold of n ($\geqq 2$) dimensions and of class C^r ($r \geqq 1$). *It can be shown that it is possible to define a Riemann metric of class C^{r-1} in M,* i. e., it is possible to define a positive definite quadratic differential form over M with coefficients which are functions of class C^{r-1} of the coördinates of the allowable coördinate neighborhoods in M.[1]

Let g be a fixed Riemann metric of class C^{r-1} in M which exists in accordance with the above result. When it is desired to measure distances independently of coördinate systems this metric may be used although within the separate coördinate systems it may sometimes be of advantage to employ the ordinary Euclidean metric.

Let f be a map of class C^r of M into the Euclidean space E_m. Let S be the set of all such maps f. We may define a metric in S in the following manner: Let ϕ and ψ be two elements of S. Put

$$d_0[\phi(x), \psi(x)] = \sqrt{[\phi^i(x) - \psi^i(x)][\phi^i(x) - \psi^i(x)]},$$
$$d_1[\phi(x), \psi(x)] = \sqrt{g^{\alpha\beta}[\phi^i{}_{,\alpha} - \psi^i{}_{,\alpha}][\phi^i{}_{,\beta} - \psi^i{}_{,\beta}]},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$d_r[\phi(x), \psi(x)] = \sqrt{g^{\alpha a}\cdots g^{\beta b}[\phi^i{}_{,\alpha\ldots\beta} - \psi^i{}_{,\alpha\ldots\beta}][\phi^i{}_{,a\ldots b} - \psi^i{}_{,a\ldots b}]},$$

where it is to be understood that Latin indices have the range $1, \cdots, m$ and that Greek indices have the range $1, \cdots, n$ and both sets of indices are to be summed when repeated in accordance with the usual convention. Note also that $d_1, \cdots, d_r$ involve successive covariant derivatives of the functions ϕ^i and ψ^i and that d_r involves the r-th covariant derivative which is the highest covariant derivative that can be formed under the hypothesis that the Riemann metric is of class C^{r-1} and the maps ϕ and ψ are of class C^r. Put

$$D[\phi(x), \psi(x)] = d_0 + d_1 + \cdots + d_r.$$

* Received October 6, 1938.

[1] H. Whitney, "Analytic coördinate systems and arcs in a manifold," *Annals of Mathematics*, vol. 38 (1937), p. 816.

Then $D[\phi(x), \psi(x)]$ is a continuous function on M. Since M is compact and separable it is bicompact and hence the function $D[\phi(x), \psi(x)]$ assumes its maximum value at a point of M. Denote this maximum value by $D(\phi, \psi)$ and define $D(\phi, \psi)$ as the distance between the points ϕ and ψ of S. Evidently $D(\phi, \psi) \geq 0$. In fact it is easily seen that with this definition of distance S is a metric space.

The map space S is complete, i. e. every Cauchy sequence in S converges. The proof is simple and will be omitted.[2]

The map or point ϕ in S will be said to be *regular* if the matrix $\| \partial\phi^i/\partial x^a \|$ has rank n at every point of M. Otherwise ϕ will be said to be *singular* and the points at which the above matrix has rank less than n will be called *singular points* of the map. The set of all such singular points will be called the *domain of singularity* of the map ϕ.

It is easily seen that *the regular maps in S form an open set.* On the basis of methods used by H. Whitney[3] the following results can be proved: (a) *if $r \geq 2$ and $m \geq 2n$ the regular maps are dense in S and* (b) *if $r \geq 2$ and $m \geq 2n + 1$ the regular topological maps form an open and dense set in S.* Since the map $f = 0$ is an element of S it follows that for $r \geq 2$ and $m \geq 2n$ the set of regular maps in S is non-vacuous, and for $r \geq 2$ and $m \geq 2n + 1$ the set of regular topological maps is non-vacuous in S.

2. In the following we take $m = n$. Denote by $\bar{U}$ the cube $-1 \leq x^a \leq 1$ in E_n and its interior, namely $-1 < x^a < 1$, by U. Similarly let $\bar{U}'$ be the cube $-\frac{1}{2} \leq x^a \leq \frac{1}{2}$ with interior $-\frac{1}{2} < x^a < \frac{1}{2}$ denoted by U'. It is possible to construct a function $H(x', \cdots, x^n)$ of class C^∞ in E_n possessing the following properties:

(a) $H = 1$ in $\bar{U}'$,
(b) $H = 0$ in $E_n - U$,
(c) $0 < H < 1$ in $U - \bar{U}'$,
(d) H is analytic in $U - \bar{U}'$.

On the basis of these properties the following lemma can be proved.

LEMMA 1. *The function H can not be homogeneous of degree -1 in an open set in U.*

---

[2] A complete metric map space has been used in purely topological embedding problems by W. Hurewicz, "Über Abbildungen von endlichdimensionalen Raümen auf Teilmengen cartesischer Raüme," *Preuss. Akad. Wissenschaften*, Berlin, vol. 24 (1933), p. 754.

[3] H. Whitney, "Differentiable manifolds," *Annals of Mathematics*, vol. 37 (1936), p. 645.

Proof. Assume H homogeneous of degree -1 in an open set $V \subset U$. Then V can not lie entirely in $\bar{U}'$ since $H = 1$ in $\bar{U}'$ by (a). Hence V must contain a point in $U - \bar{U}'$ and hence an open set W such that $W \subset V$ and $W \subset U - \bar{U}'$. By assumption H is then homogeneous of degree -1 in W. Let $x^\alpha = \xi^\alpha t$ with the ξ's constant and $t \geq 0$ define a straight line l issuing from the origin in E_n and passing through a point of W. Then by our assumption $H = c/t$ along l in W where c is a suitably chosen constant. From (c) we must have $c > 0$. Since H is analytic in $U - \bar{U}'$ by (d) it follows that H must be represented by c/t along the portion of l contained in $U - \bar{U}'$. From this fact and the continuity of H in E_n it follows that $H \neq 0$ at the intersection of l with the boundary of U. As this is in contradiction with (b) the proof is complete.

3. Let $\phi^1, \cdots, \phi^n$ be a set of n functions of class C^r $(r \geq 1)$ in U. Consider the functions $\psi^\alpha = \phi^\alpha + Hx^\beta z_\beta{}^\alpha$ in U where the z's are arbitrary constants. Also consider the functional determinant $\Delta = | \partial\psi^\alpha/\partial x^\gamma |$ in U. Let J denote the set of points in U at which Δ vanishes identically in the $z_\beta{}^\alpha$.

LEMMA 2. *The set J is nowhere dense in U.*

Proof. We have

$$\Delta = \left| \left(H\delta_\gamma{}^\beta + x^\beta \frac{\partial H}{\partial x^\gamma}\right) z_\beta{}^\alpha + \frac{\partial \phi^\alpha}{\partial x^\gamma} \right| .$$

If at a point $P \subset U$ the determinant Δ' of the coefficients of the z's in the above representation of Δ is different from zero we can choose the z's so that at P the elements of Δ can have arbitrary values, for example the values $\delta_\gamma{}^\alpha$. At such a point P the determinant Δ will not vanish identically in the z's. Denote by J' the set of points in U at which Δ' is zero. We have

$$\Delta' = \left| H\delta_\gamma{}^\beta + x^\beta \frac{\partial H}{\partial x^\gamma} \right| = H^{n-1}\left(H + x^\alpha \frac{\partial H}{\partial x^\alpha}\right).$$

Since $H \neq 0$ in U it follows that $\Delta' = 0$ if, and only if, the above expression in the parenthesis is equal to zero. By Lemma 1 this expression can not vanish in an open set $W \subset U$ since this is the condition for H to be homogeneous of degree -1 in W. Hence Δ' can not vanish in an open set in U. Any open set $W \subset U$ must therefore contain a point Q at which $\Delta' \neq 0$ and, since Δ' is a continuous function, there must be a neighborhood $W(Q) \subset W$ which contains no point of J', i. e. J' is nowhere dense in U. Hence J is nowhere dense in U since $J \subset J'$.

4. Let R be the set of rational points in U, i. e. the points in U having

rational coördinates. Consider the intersection $R \cdot J$ of R with the nowhere dense set J. Then $K = R - R \cdot J$ is dense in U and countably infinite. Let $P_1, P_2, \cdots$ denote the points of K.

To indicate the dependence of the determinant Δ on the selection of the constants $z_\beta{}^a$ let us write $\Delta = \Delta(z_\beta{}^a)$. Let O be any open set in E_n. Since $\Delta(z_\beta{}^a) \neq 0$ in the z's at P_1 we can find a set of points $z_{\beta/1} \subset O$ such that $\Delta(z^a{}_{\beta/1}) \neq 0$ at P_1. Hence there exist open sets $W(z_{\beta/1})$ with enclosures $\bar{W}(z_{\beta/1}) \subset O$ such that $\Delta(z_\beta{}^a) \neq 0$ at P_1 for $z_\beta \subset \bar{W}(z_{\beta/1})$. Similarly we can find points $z_{\beta/2} \subset W(z_{\beta/1})$ such that $\Delta(z^a{}_{\beta/2}) \neq 0$ at P_2 and then we can find open sets $W(z_{\beta/2})$ with enclosures $\bar{W}(z_{\beta/2}) \subset W(z_{\beta/1})$ such that $\Delta(z_\beta{}^a) \neq 0$ at P_2 for $z_\beta \subset \bar{W}(z_{\beta/2})$. Hence $\Delta(z_\beta{}^a) \neq 0$ at P_1 and P_2 for $z_\beta \subset \bar{W}(z_{\beta/2})$. Continuing we obtain after k steps the sets $\bar{W}(z_{\beta/1}) \supset \bar{W}(z_{\beta/2}) \supset \cdots \supset \bar{W}(z_{\beta/k})$ with $\Delta(z_\beta{}^a) \neq 0$ at $P_1, P_2, \cdots, P_k$ for $z_\beta \subset \bar{W}(z_{\beta/k})$.

Now the decreasing sequence

$$\bar{W}(z_{\beta/1}) \supset W(z_{\beta/2}) \supset \cdots \supset \bar{W}(z_{\beta/k}) \supset \cdots$$

of non-vacuous closed sets has a non-vacuous intersection (Cantor's intersection theorem). Let $Y_\beta \subset O$ be a point of this intersection. It follows that $\Delta(Y_\beta{}^a)$ does not vanish in U at any point of the set K. From the continuity of the function $\Delta(Y_\beta{}^a)$ and the fact that K is dense in U we see that $\Delta(Y_\beta{}^a)$ is different from zero over a set $T \supset K$ which is dense and open in U. Hence the complement L of T in U over which $\Delta(Y_\beta{}^a) = 0$ is nowhere dense and is evidently closed. We state these results in the following

LEMMA 3. *It is possible to choose points $Y_1, \cdots, Y_n$ in any open set in E_n such that the functional determinant $\Delta(Y_\beta{}^a)$ will vanish in U only over a set which is closed and nowhere dense in U.*

5. Let P be any point of M and denote by $N(P)$ a coördinate neighborhood of P. We suppose the coördinates x^a so chosen in $N(P)$ that $x^a = 0$ at P. By a coördinate transformation in $N(P)$ of the form $\bar{x}^a = ax^a$ ($a =$ constant) we can so enlarge (if necessary) the coördinate representation of this neighborhood that it will contain the above cube U. We now consider a covering of M by the open sets U. Since M is bicompact this covering contains a finite covering. Let such a finite covering be denoted by $U_1, U_2, \cdots, U_t$.

Now consider the metric space S (§ 1) the elements of which are the maps of class C^r ($r \geq 1$) of M into E_n. Denote by Σ the set of all points of S which correspond to maps whose domain of singularity is nowhere dense in M. We shall show that Σ is dense in S, i. e. if ϕ is any point of S and ρ any positive constant there exists a point $\psi \subset \Sigma$ such that $D(\phi, \psi) < \rho$. Put

$$\begin{cases} \phi_1{}^a = \phi^a + Hx^\beta z_\beta{}^a & \text{in } \bar{U}_1, \\ \phi_1{}^a = \phi^a & \text{in } M - \bar{U}_1, \end{cases}$$

where the z's are constants. Then ϕ_1 is a map of class C^r of M into E_n and hence is an element of S. Restrict the constants z to be the coördinates of points z_β in the open set $| x^a | < \eta$ in E_n. By Lemma 3 we can furthermore chose the $z_\beta{}^a$ so that the set of singular points of ϕ_1 in U be nowhere dense in U. In addition by taking the above constant η sufficiently small the map ϕ_1 will approximate the map ϕ as closely as desired. Hence we can secure the condition $D(\phi, \phi_1) < \rho/t$.

We now approximate the map ϕ_1 in a similar manner by a map ϕ_2 using the neighborhood U_2 for this purpose. In fact we put

$$\begin{cases} \phi_2{}^a = \phi_1{}^a + Hx^\beta z_\beta{}^a & \text{in } \bar{U}_2, \\ \phi_2{}^a = \phi_1{}^a & \text{in } M - \bar{U}_2, \end{cases}$$

and choose the constants $z_\beta{}^a$ so that the set of singular points of ϕ_2 in U_2 will be nowhere dense in U_2 and also so that $D(\phi_1, \phi_2) < \rho/t$. Then the set of singular points of ϕ_2 in $U_1 + U_2$ will be nowhere dense in $U_1 + U_2$. Continuing we obtain maps $\phi_1, \cdots, \phi_t$ such that for any k with $1 \leq k \leq t$ we have (a) $\phi_k \subset S$, (b) $D(\phi_{k-1}, \phi_k) < \rho/t$ with $\phi_0 = \phi$ and (c) the set of singular points of ϕ_k in $U_1 + \cdots + U_k$ is nowhere dense in $U_1 + \cdots + U_k$. Then $\psi = \phi_t$ is the desired map. In fact the domain of singularity of ψ is nowhere dense in M and

$$D(\phi, \psi) \leq D(\phi, \phi_1) + D(\phi_1, \phi_2) + \cdots + D(\phi_{t-1}, \psi) < t(\rho/t) = \rho.$$

We have now proved the following theorem.

THEOREM. *Let M be a compact and separable manifold of n (≥ 2) dimensions and of class C^r $(r \geq 1)$ and S the metric space (§ 1) whose elements are the maps of class C^r of M into E_n. Let Σ be the set of points in S which correspond to maps whose domain of singularity is nowhere dense in M. Then the set Σ is dense in S.*

6. Since $f = 0$ is a map of class C^r of M into E_n the set Σ is non-vacuous and the above theorem admits the following

COROLLARY I. *Any compact and separable manifold M of n (≥ 2) dimensions and class C^r $(r \geq 1)$ can be embedded in the Euclidean space of n dimensions by a map of class C^r whose domain of singularity is closed and nowhere dense in M.*

If we extend the strict definition of the Riemann space to include spaces with metric defined by a *non-negative* quadratic differential form (i. e. if we omit the restriction that this form be non-singular at every point of the space) a second corollary can be stated.

COROLLARY II. *In any compact and separable manifold M of n (≥ 2) dimensions and class C^r $(r \geq 1)$ a Riemann metric of class C^{r-1} can be defined which is non-singular and locally flat in the ordinary sense except over a closed and nowhere dense set in the space.*

UNIVERSITY OF CALIFORNIA,
LOS ANGELES.

THE DISTRIBUTION OF THE MAXIMA OF A RANDOM CURVE.*

By S. O. Rice.

Introduction. The equation

$$y = F(a_1, a_2, \cdots, a_n, x) \tag{1}$$

defines y as a function of x when the values of the parameters are assigned. We assume that y is single valued and real for a certain range of values of x and that the parameters $a_1, a_2, \cdots, a_n$ are chance variables whose distribution functions are known. When a set of parameters is selected at random y may be plotted as a function of x. We shall call a curve obtained in this way a random curve.

In general, a random curve will have a number of maxima, their number and positions depending upon the particular values of $a_1, \cdots, a_n$ which happened to be drawn. Here the distribution of the maxima of such random curves is studied. Although this problem is of some physical interest I have been unable to find references to any earlier work. Problems of this nature occur in the investigation of the current reflected by small random irregularities along telephone transmission lines.

We shall speak of the probability $P(R)$ that a maximum of a random curve will occur in a given region R of the (x, y) plane and also of the expected number of maxima $E(R)$ in that region. To make these terms clear, we suppose N sets of the parameters $(a_1, \cdots, a_n)$ are drawn and the corresponding N curves plotted. Let N_c be the number of curves which have one or more maxima in R and N_m the total number of maxima in R. Thus, if the first curve has one maximum in R, the second curve none, the third two, and so on until the last one has, say, three maxima, then

$$N_c = 1 + 0 + 1 + \cdots + 1 \quad \text{and} \quad N_m = 1 + 0 + 2 + \cdots + 3.$$

$P(R)$ and $E(R)$ are then given by

$$P(R) = \lim_{N\to\infty} \frac{N_c}{N}, \qquad E(R) = \lim_{N\to\infty} \frac{N_m}{N}$$

where the limits are assumed to exist.

The discussion given in the section entitled "Theory" establishes the

* Received July 15, 1938.

following result: If suitable conditions are satisfied the probability that the random curve

$$y = F(a_1, a_2, \cdots, a_n, x) \tag{2}$$

has a maximum in the rectangle $(x_0, x_0 + dx_0;\ y_0, y_0 + dy_0)$, dx_0 and dy_0 being of the same order of magnitude, is $p(x_0, y_0)\,dx_0 dy_0$ where

$$p(x_0, y_0) = -\int_{-\infty}^{0} P(y_0, 0, \zeta)\zeta d\zeta. \tag{3}$$

$P(\xi, \eta, \zeta)\Delta\xi\Delta\eta\Delta\zeta$ is the probability to within first order terms that, when $x = x_0$, $F = F(a_1, a_2, \cdots, a_n, x)$ and its first two partial derivatives with respect to x lie within the intervals

$$\begin{aligned} \xi &< F < \xi + \Delta\xi \\ \eta &< \frac{\partial F}{\partial x} < \eta + \Delta\eta \\ \zeta &< \frac{\partial^2 F}{\partial x^2} < \zeta + \Delta\zeta. \end{aligned}$$

$P(\xi, \eta, \zeta)$ may be determined from $F(a\ , \cdots, a_n, x)$ and the distribution functions of the chance variables $a_1, a_2, \cdots, a_n$. In general, it will involve x_0 as a parameter. The probability that a maximum occurs in the strip $(x_0, x_0 + dx_0)$ is

$$dx_0 \int_{-\infty}^{+\infty} p(x_0, y)\,dy.$$

When $F(a_1, \cdots, a_n, x)$ and $P(\xi, \eta, \zeta)$ satisfy the conditions assumed in the next section the probability of the occurrence of a maximum in the rectangles and strips just mentioned is equal to the expected number of maxima in the respective regions. However, for larger regions this is no longer true, and

$$\int_a^b dx \int_c^d dy p(x, y)$$

is not the probability that a maximum of the curve will occur in the rectangle $(a, b;\ c, d)$, but is the expected number of maxima in that region. Incidentally, considerations of this sort enable us to form an estimate of the sinuosity of a random curve.

Similar results hold for the distribution of the minima of a random curve, the expression for the probability that a minimum occurs in the rectangle $(x_0, x_0 + dx_0;\ y_0, y_0 + dy_0)$ being

$$dx_0 dy_0 \int_0^{\infty} P(y_0, 0, \zeta)\zeta d\zeta.$$

The example which is used to illustrate the theory indicates that the restrictions placed on $F(a_1, \cdots, a_n, x)$ in the derivation of equation (3) are harsher than necessary.

Theory. We assume that $F(a_1, \cdots, a_n, x)$ and its first three derivatives with respect to x are continuous, single-valued functions of x, and are bounded, but the bounds are not zero, for all possible values of the parameters and for the values of x in the range of interest. It is also assumed that the distribution functions of the parameters and the form of $F(a_1, \cdots, a_n, x)$ are such that the distribution function $P(\xi, \eta, \zeta)$ is a continuous function of its arguments for any fixed value of x_0. For convenience we shall write $F(a_1, \cdots, a_n, x)$ as $F(x)$ and indicate derivatives with respect to x by primes.

In the proof of equation (3) we find a set of conditions, denoted by I, which a random curve satisfies if it has a maximum in the rectangle $(x_0, x_0 + \Delta x;\, y_0, y_0 + \Delta y)$. We shall call this rectangle the elementary rectangle, and take Δx and Δy to be of the same order of magnitude. However, a random curve may not have a maximum in the elementary rectangle yet still satisfy conditions I. Another set of conditions II are then found which when satisfied by a random curve guarantee that it has a maximum in the rectangle. However, a random curve may have a maximum in the rectangle yet not satisfy conditions II.

The probability $q(\Delta x, \Delta y)$ that a random curve has one or more maxima [here we wish to emphasize the exact meaning of $q(\Delta x, \Delta y)$ so we say "one or more maxima" instead of merely "a maximum"] in the elementary rectangle is less than the probability that conditions I are satisfied and greater than the probability that conditions II are satisfied. It is shown that when Δx and Δy approach zero the expressions for these last two probabilities approach the same limiting form. Hence $q(\Delta x, \Delta y)$ also approaches the same form in the limit. The expression is found to be $p(x_0, y_0)\Delta x \Delta y$ where $p(x_0, y_0)$ is given by equation (3). Finally it is shown, under the assumptions stated above, that as the rectangle becomes small the probability of two or more maxima occurring in it becomes small in comparison with the probability that only one occurs.

If a random curve $y = F(x)$ has a maximum in the elementary rectangle it must satisfy each one of the three inequalities

$$\mathrm{I}\begin{cases} (4) \qquad y_0 - M_2(\Delta x)^2 \leq F(x_0) \leq y_0 + \Delta y + M_2(\Delta x)^2 \\ (5) \qquad -M_3(\Delta x)^2 < F'(x_0) \\ (6) \qquad F'(x_0) + \Delta x F''(x_0) < 2M_3(\Delta x)^2 \end{cases}$$

where M_2 and M_3 are the upper bounds, for all allowable values of the para-

meters and x, of $|F''(x)|$ and $|F'''(x)|$. These are the conditions referred to above as conditions I. Let the maximum be at the point (x_1, y_1). Then $y_1 = F(x_1)$, $F'(x_1) = 0$, $x_0 < x_1 < x_0 + \Delta x$, and

$$y_0 < F(x_1) < y_0 + \Delta y. \tag{7}$$

If $x_0 \leq x \leq x_0 + \Delta x$, the mean value theorem gives

$$|F'(x)| = |F'(x) - F'(x_1)| < M_2 \Delta x \tag{8}$$

$$|F(x) - F(x_0)| \leq \Delta x \operatorname{Max} |F'(x)| < M_2 (\Delta x)^2. \tag{9}$$

Setting $x = x_1$ in the last result and using (7) shows that the inequalities (4) are satisfied by any random curve having a maximum in the elementary rectangle.

If $F'(x_0) \geq 0$ condition (5) is automatically fulfilled. If $F'(x_0) < 0$ the random curve must have a minimum before it has the assumed maximum. Thus there are two points in the interval $x_0 < x < x_0 + \Delta x$ where $F'(x) = 0$. By Rolle's theorem there is also a point, say $x = x_2$, such that $F''(x_2) = 0$. The same reasoning used to establish (9) may be used to show that

$$|F'(x_0)| = |F'(x_1) - F'(x_0)| < M_3 (\Delta x)^2. \tag{10}$$

Thus (5) is also satisfied if $F'(x_0) < 0$.

In the same way it may be shown that

$$F'(x_0 + \Delta x) < M_3 (\Delta x)^2. \tag{11}$$

From Lagrange's form for the remainder in Taylor's series,

$$F'(x_0 + \Delta x) = F'(x_0) + \Delta x F''(x_0) + \frac{(\Delta x)^2}{2} F'''(x_0 + \theta \Delta x), \tag{12}$$

where $0 \leq \theta \leq 1$, it follows that

$$F'(x_0 + \Delta x) > F'(x_0) + \Delta x F''(x_0) - M_3 (\Delta x)^2,$$

and this together with inequality (11) gives inequality (6).

If a random curve $y = F(x)$ satisfies the inequalities

$$\text{II} \begin{cases} (13) & y_0 + M_2(\Delta x)^2 \leq F(x_0) \leq y_0 + \Delta y - M_2(\Delta x)^2 \\ (14) & 0 < F'(x_0) \\ (15) & F'(x_0) + \Delta x F''(x_0) < -M_3(\Delta x)^2 \end{cases}$$

then it has at least one maximum in the elementary rectangle. These conditions were referred to as conditions II. It follows from equation (12) that

$$F'(x_0 + \Delta x) < F'(x_0) + \Delta x F''(x_0) + M_3(\Delta x)^2$$

which together with (15) shows that $F'(x_0 + \Delta x) < 0$. From this and (14) it follows that there is at least one maximum of $y = F(x)$ in the interval $x_0 < x < x_0 + \Delta x$, say at $x = x_1$, so $F'(x_1) = 0$. Inequality (9) then holds and it, together with (13), shows that

$$y_0 < F(x) < y_0 + \Delta y.$$

Thus the maxima of the random curve lie within the elementary rectangle when $x_0 < x < x_0 + \Delta x$.

It was pointed out at the beginning of this section that the probability of conditions I being satisfied is greater than the probability $q(\Delta x, \Delta y)$ of a random curve having a maximum in the elementary rectangle, which in turn exceeds the probability of conditions II being satisfied. From the definition of the distribution function $P(\xi, \eta, \zeta)$ it follows that the probability of conditions I being satisfied is

$$\int_{y_0 - M_2(\Delta x)^2}^{y_0+\Delta y+M_2(\Delta x)^2} d\xi \int_{-M_3(\Delta x)^2}^{M_1} d\eta \int_{-M_2}^{-\eta/\Delta x + 2M_3\Delta x} d\zeta P(\xi, \eta, \zeta)$$

where M_1 is the upper bound of $|F'(x)|$. The limits of integration for ξ, η and ζ are fixed by conditions (4), (5) and (6) respectively, together with the conditions $\eta < M_1$ and $\zeta > -M_2$. However, as Δx becomes small the effective upper limit of η becomes $M_2\Delta x + 2M_3(\Delta x)^2$ rather than M_1 because values of η larger than the former make the upper limit of integration of ζ less than $-M_2$; and $P(\xi, \eta, \zeta)$ is zero for these values.

Since $P(\xi, \eta, \zeta)$ is continuous, it is certainly finite, and it may be seen that the integral just above differs from

$$\int_{y_0}^{y_0+\Delta y} d\xi \int_0^{M_2\Delta x} d\eta \int_{-M_2}^{-\eta/\Delta x} d\zeta P(\xi, \eta, \zeta),$$

which is of order $\Delta x \Delta y$, by terms of order $\Delta y(\Delta x)^2$. The same line of reasoning shows that the probability of conditions II being satisfied also differs from this integral by terms of order $\Delta y(\Delta x)^2$. Therefore, this integral can not differ from $q(\Delta x, \Delta y)$ by more than terms of order $\Delta y(\Delta x)^2$. By changing the order of integration, using the mean value theorem for integrals and the fact that $P(\xi, \eta, \zeta)$ is a continuous function of all its variables, it may be shown that

$$\lim_{\Delta x, \Delta y \to 0} \frac{q(\Delta x, \Delta y)}{\Delta x \Delta y} = -\int_{-M_2}^{0} P(y_0, 0, \zeta)\zeta d\zeta.$$

Since $P(y_0, 0, \zeta)$ is zero for ζ less than $-M_2$, the lower limit of integration

may be replaced by $-\infty$ and we obtain equation (3). This shows that when the conditions mentioned in the first paragraph of this section are fulfilled $p(x_0, y_0)\Delta x\Delta y$ is the probability that one or more maxima will lie within the elementary rectangle.

However our hypothesis allows us to go further and show that this probability differs from the probability that only a single maximum will occur in the rectangle by terms of order higher than $\Delta x\Delta y$. If a random curve has two or more maxima in the elementary rectangle equation (4) of conditions I is still satisfied. Equation (5) may be replaced by $|F'(x_0)| < M_3(\Delta x)^2$. This follows from equation (10) since there are at least two points where $F'(x) = 0$. Also equation (6) may be replaced by $|F''(x_0)| < M_3\Delta x$, which is similar to (8). When these conditions are used to fix the limits of integration it is seen that the probability of a random curve having two or more maxima in the rectangle is not greater than

$$\int_{y_0-M_2(\Delta x)^2}^{y_0+\Delta y+M_2(\Delta x)^2} d\xi \int_{-M_2(\Delta x)^2}^{M_3(\Delta x)^2} d\eta \int_{-M_3\Delta x}^{M_3\Delta x} d\zeta P(\xi, \eta, \zeta).$$

Since this integral is of order $\Delta y(\Delta x)^3$ and since the probability of the occurrence of one or more maxima is of order $\Delta y\Delta x$, it follows that as the rectangle becomes small the probability of two or more maxima occurring in it becomes negligible in comparison with the probability that only one occurs.

Application of theory to an example. It would be interesting to apply the theory to determine the distribution of the maxima of a curve defined by

$$y = a_1 \sin x + a_2 \sin 2x + \cdots + a_n \sin nx$$

which is similar to the case occurring in the telephone transmission line problem. However, when this done it is found that the mathematical details become rather involved despite the straightforwardness of the work. Since this reduces its value as an example we consider instead one of the simplest problems; namely, that of studying the distribution of the maxima of curves of the form

$$y = a_0 + 2a_1x + a_2x^2$$

where the parameters a_0, a_1, a_2 are independent, each being distributed about zero according to a normal law with unit standard deviation. The requirements that y and its first two derivatives be bounded are not met by this example, but the reasonableness of the results indicate that the answers are correct.

When we differentiate and obtain the first and second derivatives of y,

$$y' = 2a_1 + 2a_2 x$$
$$y'' = 2a_2,$$

we see that y has a maximum only if a_2 is less than zero, the probability of which is one-half. Also, if the maximum occurs it is at the point $x = -a_1/a_2$ and the distribution function of $-a_1/a_2$ accordingly, gives us the distribution of the maximum of y. These two conclusions will be used to check the theory.

In order to apply the equation

$$(3) \qquad p(x,y) = -\int_{-\infty}^{0} P(y, 0, \zeta)\zeta d\zeta$$

to find the probability $p(x,y)dxdy$ that a maximum will occur in the rectangle $(x, x+dx; y, y+dy)$ it is necessary first to determine the distribution function $P(y, y', y'')$. The expressions for y and its first two derivatives show that they are linear functions of the normally distributed chance variables a_0, a_1, a_2. We may therefore apply a theorem [1] in probability theory to find the required distribution functions.

$$P(y, y', y'') = \frac{(2\pi)^{-3/2}}{\sqrt{B}} \exp\left[-\frac{1}{2B}(B_{11}y^2 + B_{22}y'^2 + B_{33}y''^2 + 2B_{12}yy' + 2B_{13}yy'' + 2B_{23}y'y'')\right]$$

where B is the determinant

$$\begin{vmatrix} 1 + 4x^2 + x^4 & 4x + 2x^3 & 2x^2 \\ 4x + 2x^3 & 4 + 4x^2 & 4x \\ 2x^2 & 4x & 4 \end{vmatrix} = 16$$

and B_{ij} is the cofactor of the element in the i-th row and j-th column. Since in equation (3) y' is zero we are interested only in B_{11}, B_{13}, and B_{33}.

$$B_{11} = 16, \qquad B_{13} = 8x^2, \qquad B_{33} = 4(1 + x^2 + x^4),$$

and we have

$$P(y, 0, y'') = \frac{(2\pi)^{-3/2}}{4} \exp\left[-\frac{y^2}{2} - \frac{x^2yy''}{2} - \frac{1 + x^2 + x^4}{8} y''^2\right].$$

The integral for the distribution function $p(x,y)$ of the maxima becomes

$$p(x,y) = -\frac{(2\pi)^{-3/2}}{4}\int_{-\infty}^{0} \zeta d\zeta \exp\left[-\frac{y^2}{2} - \frac{x^2y\zeta}{2} - \frac{1 + x^2 + x^4}{8}\zeta^2\right]$$
$$= \frac{2s^2}{(2\pi)^{3/2}} e^{-y^2/2} + \frac{s^3x^2y}{\pi\sqrt{2}} e^{-(1+x^2)y^2s^2}[1 + erf(x^2sy)]$$

[1] A discussion of the general theorem has been given by van Uven, *K. Akad. v. Wetensch. Amsterdam Proc.* 16, 1124-35 (1914); see also S. O. Rice, *Quarterly Journal of Mathematics*, Oxford Series, vol. 9 (1938), pp. 1-4.

where

$$s = 1/\sqrt{2(1 + x^2 + x^4)}, \qquad erf(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-t^2} dt.$$

The probability that a maximum lies in the strip between x and $x + dx$ is obtained by integrating $p(x, y)$ from $y = -\infty$ to $y = +\infty$ and multiplying by dx. The integration is not difficult to perform and gives

$$dx \int_{-\infty}^{+\infty} p(x, y) dy = \frac{dx}{2\pi(1 + x^2)}. \tag{16}$$

We are now in a position to check the conclusions we drew at the beginning of this section. The expected number of maxima per curve is obtained by integrating the probability that a maximum occurs in a strip of width dx from $x = -\infty$ to $x = +\infty$. Thus the expected number of maxima in the interval $-\infty < x < +\infty$ for the quadratic curves of our example is

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{\infty} dy p(x, y) = \int_{-\infty}^{+\infty} \frac{dx}{2\pi(1 + x^2)} = \tfrac{1}{2}$$

which agrees with our first conclusion.

To use the second conclusion to check our theory we have to show that the point $x = -a_1/a_2$ at which the maximum occurs has a distribution function which is closely related to the expression (16) which our theory gives for the probability that a maximum lies in the strip $(x, x + dx)$. The exact relation which must be satisfied is as follows:

$$\begin{aligned}&[\text{Probability that a max. lies in } x, x + dx] \\ &\quad = [\text{Probability that } a_2 < 0][\text{Probability that } x < -\frac{a_1}{a_2} < x + dx].\end{aligned}$$

Equation (16) tells us that the expression on the left is $dx/[2\pi(1 + x^2)]$. The probability that a_2 is negative is $1/2$. Furthermore, it is not difficult to verify that the distribution function of $x = -a_1/a_2$ where a_1 and a_2 are normally distributed chance variables is $1/\pi(1 + x^2)$. The two sides of the above equation are therefore equal and the second conclusion checks our theory.

Bell Telephone Laboratories.

THE CLAMPED SQUARE SHEET.* [1]

By D. G. Bourgin.

The equations of elasticity for the thin plate problem contain planar tension and shear terms, together with bending and vertical shear stresses. For the moderately thin plate with deflections small compared with the thickness the planar stresses and second order terms in the strains may be neglected, thus leaving a single thin plate linear equation given in the standard treatises and extensively investigated in the literature.

The term "sheet" is introduced in this paper to designate the *opposite* extreme, namely that of ultra-thin plates with large vertical deflections for which the plane stresses alone are significant. This sort of problem arises in aeronautics and kindred fields where paper-thin coverings are used. Second order terms must be retained. Two non-linear equations are fundamental. The superposition principle in the deflections is no longer applicable. Although an Airy's type function is introduced, it refers to plane stresses and not to vertical deflections as in the case of plates. New boundary conditions appear and require special interpretation both in the discussion of the solution and in the uniqueness demonstration. Other differences with the moderately thin plate investigations will be apparent in the sequel.

The specific concern of the present work is the clamped square sheet. So far as the writer is aware, this problem has been treated by two authors only; namely, Hencky [2] and Foppl.[3] These authors are avowedly interested solely in numerical approximations and make no pretense of attempting a rigorous solution.

The method of Hencky consists in replacing derivatives by finite differences and solving the resulting system of simultaneous linear algebraic equations. Even for three subdivisions the labor of computation is quite formidable. Moreover, to the writer's mind, the boundary conditions used are incorrect. The method of Foppl consists in approximating the displacements z, u, and v by single constant products of the form $c \sin x/a \sin y/a$; $u_0 \sin 2x/a \sin y/a$; $v_0 \sin x/a \sin 2y/a$, (c, u_0, v_0 are undetermined coefficients) plus the further

* Received September 8, 1937; Revised July 5, 1938.

[1] Presented to the American Mathematical Society, January 1, 1936.

[2] H. Hencky, *Z. für Ang. Math. und Mech.*, Bd. 1 (1914), pp. 81, 423.

[3] A. and L. Foppl, *Drang und Zwang*, 2nd ed., p. 226. This joint work will be referred to in the singular.

arbitrary assumption that the pressure is then constant, and then applying the currently popular Ritz method to determine the relations between the three constants. The objections to be urged against his work are that the three assumed approximations are not compatible with the proper boundary conditions, and no criteria are provided for determining the deviation of any results obtained from the correct expressions.

The problem formulated by Hencky and Foppl is that of uniform pressure. The determination of the stresses and strains for assigned pressure will generally involve long and unwieldy expressions. In principle, the problem of uniform pressure may perhaps be solved by using series expansion for the vertical deflection, the coefficients being determined to satisfy the uniform pressure condition. The general method is clear from the case, treated here, of a single such term in the vertical deflection. We may alternatively look upon the present research as a *complete* solution of a sort of converse problem—namely that of the determination of the stresses, strains and pressure for a particular assigned vertical deflection. For comparison with Foppl's analysis, the deflection z is, except for notation, precisely that adopted by him. It is thus possible, as a result of our work, to gauge the departure from the constancy that Foppl assumes for the pressure. Moreover, the accurate expressions for u and v may be used to determine the adequacy of the arbitrary approximations made by him.

In connection with the practical computations an important paper by March [4] on the moderately thin plate ought be cited. His procedure, like ours, involves the solution of an infinite system of linear algebraic equations though his empirical method of numerical solution is given no justification in his paper. He demonstrates uniqueness of the Neumann expansion, our Eq. 33, and anticipates the idea behind our Eq. 38.

In the main, two methods are followed below. The first makes the stresses central, which seems to be a comparative novelty in Elasticity theory; and the second, the strains. Uniqueness of the solutions is established, and the procedure is rigorously founded.

We make our problem concrete by postulating a square sheet of side π placed along the coördinate axes in the first quadrant. For convenience in stating the formulae the following conventions will be observed:

P before an expression containing x and y indicates that a similar expression with y, x replacing x, y is to be added. Q implies that the second expression is to be subtracted.

[4] H. W. March, *Transactions of the American Mathematical Society*, vol. 27 (1925), p. 307. The writer is indebted to the referee for this and other references which have allowed elision of overlapping material.

S, C, T, Sh, Ch, Th, Sch, are abbreviations for the sine, cos, tan, sinh, cosh, tanh, sech, respectively. Example:

$$\begin{matrix} P \\ Q \end{matrix}\ Shx\,Cy = \begin{matrix} \sinh x \cos y + \sinh y \cos x \\ \sinh x \cos y - \sinh y \cos x. \end{matrix}$$

The differential system. The following equations are fundamental.[5]

$$\begin{aligned} &\epsilon_1 = u_x + z_x^2/2 = \partial u/\partial x + (\partial z/\partial x)^2/2 \\ (1)\qquad &\epsilon_2 = v_y + z_y^2/2 \\ &\epsilon_{12} = v_x + u_y + z_x z_y \end{aligned}$$

$$\begin{aligned} &T_1 = \lambda\Delta + 2\mu\epsilon_1 = A(\epsilon_1 + \sigma\epsilon_2), \qquad \Delta = \epsilon_1 + \epsilon_2 \\ (1.1)\qquad &T_2 = \lambda\Delta + 2\mu\epsilon_2 = A(\epsilon_2 + \sigma\epsilon_1) \\ &S_1 = \mu\epsilon_{12} = A(1-\sigma)\epsilon_{12}/2, \qquad A = \lambda + 2\mu = \epsilon/1 - \sigma^2 \end{aligned}$$

where λ, μ are the usual elastic constants, σ is Poisson's ratio, the ϵ's are the strains and the T_i's are the tensional stresses with S_1 the planar shear. u, v, z are the deflections in the x, y, z direction, respectively.[6]

The assumption that the stresses are derivable from an Airy's function is sufficient to uniquely determine the correct equations for the sheet. Thus with

$$(2)\qquad T_1 = U_{yy}, \qquad T_2 = U_{xx}, \qquad S_1 = -U_{xy}$$

there follows

$$(3)\qquad \frac{\partial T_1}{\partial x} + \frac{\partial S_1}{\partial y} = 0$$

$$(3.1)\qquad \lambda\frac{\partial\Delta}{\partial x} + 2\mu\frac{\partial\epsilon_1}{\partial x} + \mu\frac{\partial\epsilon_{12}}{\partial y} = (\lambda+\mu)\frac{\partial\Delta}{\partial x} + \mu\left[\nabla^2 u + \frac{\partial z}{\partial x}\nabla^2 z\right] = 0$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$. Similarly

$$(3.2)\qquad (\lambda+\mu)\frac{\partial\Delta}{\partial y} + \mu[\nabla^2 v + z_y\nabla^2 z] = 0$$

arises from $\partial T_2/\partial y + \partial S_1/\partial x = 0$. On differentiating equations 3.1 and 3.2 with respect to x and y there results, after evident reduction,

$$(4)\qquad (\lambda + 2\mu)\nabla^2\Delta + 2\mu(z_{xx}z_{yy} - z_{xy}^2) = 0.$$

From Eq. 1.1, $(\lambda+\mu)\Delta = T_1 + T_2 = T$

[5] A. E. H. Love, *Theory of Elasticity*, 4th ed. for notation.

[6] For very large or "finite" displacements, the far reaching theory of F. D. Murnaghan, *American Journal of Mathematics*, vol. 59 (1937), p. 235, must furnish the more general equations corresponding to Eqs. 1, 1.1.

$$\therefore \qquad \nabla^2 T = -\frac{4\mu(\lambda+\mu)}{\lambda+2\mu}\,[z_{xx}z_{yy}-z_{xy}{}^2]$$

$$(5) \qquad = -A(1-\sigma^2)\,[z_{xx}z_{yy}-z_{xy}{}^2].$$

This is in agreement with the result given by Love.[7] The equation of equilibrium for the vertical stresses is easily derived in the form [7]

$$(6) \qquad T_1 z_{xx} + T_2 z_{yy} + 2S_1 z_{xy} = -p.$$

For the clamped sheet, the conditions at $x=0,\ a\pi$ are $u,\ v,\ z=0$ and hence the y derivatives to any order vanish. A similar statement holds for the boundaries $y=0,\ a\pi$ with x replacing y as the differentiating variable. To determine the boundary conditions in terms of the stresses we revert to Eq. 1. 1 from which it is patent that

$$(7) \qquad T_2\,|_{x=0,a\pi} = A\sigma(u_x{}^2 + z_x{}^2/2)\,|_{x=0,a\pi} = \sigma T_1\,|_{x=0,a\pi} = \rho T\,|_{x=0,a\pi}$$

$$(7.1) \qquad T_1\,|_{y=0,a\pi} = \rho T\,|_{y=0,a\pi}$$

with $\rho = \sigma/\sigma+1$.

Evidently

$$(8) \qquad \partial T_1/\partial x\,|_{x=0,a\pi} = A(u_{xx} + z_x z_{xx} + \sigma v_{xy})\,|_{x=0,a\pi}$$

$$(8.1) \qquad \partial S_1/\partial y\,|_{x=0,a\pi} = (A(\sigma-1)/2)\,v_{xy}\,|_{x=0,a\pi}.$$

According to Eq. 2 we have

$$(9) \qquad (u_{xx} + z_x z_{xx})\,|_{x=0,a\pi} = -\frac{1+\sigma}{2}\,v_{xy}\,|_{x=0,a\pi}.$$

Now

$$(8.2) \qquad \partial T_2/\partial x\,|_{x=0,a\pi} = A(v_{xy} + \sigma(u_{xx} + z_x z_{xx}))\,|_{x=0,a\pi}.$$

Combining Eqs. 8 and 1. 4 we may write

$$(7.2) \quad E(x,y)\,\Big|_{x=0,a\pi} \equiv \left(\frac{\partial T}{\partial x} + l\,\frac{\partial T_2}{\partial x}\right)\Big|_{x=0,a\pi} = \left(\frac{\partial}{\partial x}\,\nabla^2 U + l U_{xxx}\right)\Big|_{x=0,a\pi} =$$

with $l = -(1+\sigma)/(2+\sigma)$. Eq. 7. 2 involves the existence of third derivatives on the boundary. The physical implications are retained in the weaker condition, suitably interpreted,

$$(7.21) \qquad \begin{aligned} &\int_{a\pi/2}^{y} E(x,y)\,|_{x=0,a\pi}\,dy = 0 \\ &\int_{a\pi/2}^{x} E(y,x)\,|_{y=0,a\pi}\,dx = 0. \end{aligned}$$

[7] Love, *loc. cit.*, p. 558.

Eq. 7. 21 must be interpreted as a limit result for a sequence of integrals taken along parallels to the boundaries. This convention is denoted by the letter F. Specifically, F implies in detail

$$(7.3) \qquad \int_{a\pi/2}^{y_0} U_{aax}(x_0, y)\,dy = \underset{\bar{x},\bar{y}\to x^0,y^0}{L}\left(\int_{a\pi/2}^{\bar{y}} U_{aax}(\bar{x}, y)\,dy\right)$$

where $\bar{x}, \bar{y}$ are the coördinates of an inner point and x^0, y^0 of a point on the boundary of the square and a may be either x or y. We observe in this connection that the continuity of $U_{aax}(x, y)$ in both variables in the interior of the square is implied in the orthodox conception of a solution of Eq. 5. Fortunately, the solution U obtained in this paper is furthermore such that the integral on the right-hand side of Eq. 7. 3 is of class C^0 in the closed square.[8] Accordingly, we are spared the otherwise necessary consideration of mode of approach of $\bar{x}, \bar{y}$ to x^0, y^0. Of course the integral on the left side of Eq. 7. 3 may also exist as an ordinary boundary integral with a value differing from that assigned by F. Convention F and the boundary conditions in the stresses, namely Eq. 7, 7. 1, (7. 2) and 7. 21, do not occur in the literature of Elasticity Theory.

Uniqueness of solution. The set of Eqs. 5, 6, 7, 7. 1 and 7. 21, involve the dependent variables U, p, z. We now demonstrate that for assigned z of class C^2 in the closed square, the solution U, supposed to exist, is unique except for linear terms in x and y under fairly general conditions. Actual solutions are exhibited later for U and p for the special assignment in Eq. 11.

We require of a solution U that (a) Eq. 5 is satisfied in the open square. (b) U_{aabb} is absolutely integrable over any closed rectangle in the open square. a, b may separately be either x or y. (c) U is of class C^3 in the open and C^2 in the closed square. (d) All indefinite integrals in the x or y direction of third derivatives of U are of class C^0 in the closed square. (e) Eqs. 7, 7. 1 and 7. 21 are satisfied. Eq. 7. 21 holds in the sense F for the boundary Γ of the square.

Suppose there could be two solutions U^a and U^b. We form $U'' = U^a - U^b$ which satisfies the continuity and boundary value conditions assumed for U^a and U^b and is a solution of Eq. 5 with right-hand side 0.

Consider the Calculus of Variations problem connected with minimizing[9]

---

[8] Throughout this paper the terminology class C^n implies continuity through derivatives of order n in both independent variables.

[9] This variational integral is introduced in an *ad hoc* fashion merely to establish uniqueness, and involves the Airy's function in the plane stresses of a sheet whose edges are clamped. There is a striking formal resemblance (when $1 + \sigma$ is replaced by $1 - \sigma$)

$$(10) \qquad J = \int_0^{a\pi} \int_0^{a\pi} \{(\nabla^2 U)^2 - 2(1+\sigma)(U_{xx}U_{yy} - U_{xy}^2)\} dxdy$$

$$(10.1) \qquad \equiv \int_0^{a\pi} \int_0^{a\pi} \{U_{xx}^2 + U_{yy}^2 - 2\sigma U_{xx}U_{yy} + 2(1+\sigma)U_{xy}^2\} dxdy$$

for functions U subject to the continuity stipulations catalogued above for U. The separate integrals in Eqs. 10 or 10. 1 exist and are finite since U is of class C^2 in the closed square. With $\sigma < 1$ the sum of the first three terms of the integrand of Eq. 10. 1 is ≥ 0. Accordingly, the integrand is positive definite in the quantities U_{xx}, U_{yy}, U_{xy}. Hence the minimum is actually attained by the unique choice,

$$(10.2) \qquad U_{xx} = U_{yy} = U_{xy} = 0$$

and the continuity requirements preclude existence of any exceptional points in the closed square. The symmetry of U as regards x and y bars the introduction of linear x and y terms obtained on integrating the second derivative in Eq. 10. 2. Thus, designating the solving function as U', we have

$$(10.3) \qquad U' = C + sx + ty.$$

The solution of the variational problem may alternatively be determined from

$$(10.4) \qquad \delta J = 0 = L_{\epsilon_i, \eta_j \to 0} \delta J_{\eta_i, \epsilon_j}$$
$$= L_{\epsilon_i, \eta_j \to 0} \left\{ \int_{\eta_1}^{a\pi - \eta_2} \int_{\epsilon_1}^{a\pi - \eta_2} \delta U L(U) dxdy + \int_{\Gamma_{\epsilon_i, \eta_j}} I_1 \frac{\partial \delta U}{\partial n} ds + \int_{\Gamma_{\epsilon_i, \eta_j}} I_2 \delta U ds \right\},^{10} \quad (i, j = 1, 2)$$

where $\Gamma_{\epsilon_i, \eta_j}$ is the boundary of a rectangle, $\epsilon_1 \leq x \leq a\pi - \epsilon_2$, $\eta_1 \leq y \leq a\pi - \eta_2$, $\epsilon_i, \eta_j > 0$ interior to the square. δU may be taken as an arbitrary function of class C^2 in the closed square. s and n refer to directions along and normal to the boundary. The application of Green's theorem implied in the right side of Eq. 10. 4, for each of the approximating rectangles, is justified in view of hypotheses (b) and (c) and the restriction on δU.[10] We proceed to show that U'' is the solution required in Eq. 10. 4.

---

to the potential energy expression in the vertical strains for the free boundary problem of the thin plate. Cf. Courant-Hilbert, *Methoden der Mathematischen Physik*, 1st ed., p. 149.

[10] For completeness we should include the conditions

$$L_{y \to \beta} U_{xy}(\alpha, y) = L_{x \to \alpha} U_{xy}(x, \beta)$$

where α, β are coördinates of a corner. However, these corner conditions are automatically fulfilled in view of (c).

It is at once verified that the Euler Lagrange equation, $L(U)=0$, is precisely the homogeneous form of Eq. 5 and is satisfied by $U=U''$. Hence the limit of the double integral in Eq. 10. 4 is 0.

Next we remark the easily verified relation

$$I_1\,|_{x=0,a\pi} \equiv [-\sigma\nabla^2 U + (1+\sigma)U_{xx}]\,|_{x=0,a\pi} \tag{10.5}$$

and the analogous identity for the sides $y=0, a\pi$. Hence, in view of Eqs. 7, 7. 1, I_1 vanishes on the boundary Γ for $U=U''$. According to (c) I_1 is continuous in the closed square. Hence the limit of the first boundary integral in Eq. 10. 4 is identical with the ordinary integral over the boundary Γ and has the value 0.

We turn to the second boundary integral. Evidently

$$\begin{aligned} I_2\,|_{x=\epsilon_1,a\pi-\epsilon_2} &\equiv E(x,y)\,|_{x=\epsilon_1,a\pi-\epsilon_2} \\ I_2\,|_{y=\eta_1,a\pi-\eta_2} &\equiv E(y,x)\,|_{y=\eta_1,a\pi-\eta_2}. \end{aligned} \tag{10.6}$$

On integrating by parts, as is permissible in view of (c) and the restriction on δU, we have for instance

$$\begin{aligned} \int_{\eta_1}^{a\pi-\eta_2} E(\epsilon,y)\delta U(\epsilon,y)\,dy = \Big[\Big(\int^y E(\epsilon,y)\,dy\Big)\delta U(\epsilon,y)\,\Big|_{\eta_1}^{a\pi-\eta_2} \\ -\int_{\eta_1}^{a\pi-\eta_2}\frac{\partial\delta U(\epsilon,y)}{\partial y}\Big(\int^y E(\epsilon,y)\,dy\Big)dy\Big]. \end{aligned} \tag{10.61}$$

In view of (d) the indefinite integral of $E(\epsilon,y)$ is of class C^0 in ϵ and y. It follows that each of the two expressions on the right-hand side are of class C^0 in ϵ and η_j for $0\leq {\epsilon \atop \eta_j} \leq a\pi$, $j=1,2$. The order of the operations of integration and passage to the limit $\epsilon\to 0$ for the definite integral on the right side of Eq. 10. 61 may therefore be interchanged. We recall the interpretation of Eq. 7. 21 in the sense F. It is clear then that, with $U=U''$, the limit of the right-hand side of Eq. 10. 61, for $\epsilon, \eta_j \to 0$, $j=1,2$, is 0 and is taken on uniformly with respect to ϵ and η_1, η_2. An analogous result holds for approach to the sides $x=a\pi$ or $y=0, a\pi$ of the appropriate integrals of the type of the left side of Eq. 10. 61. Accordingly, the second boundary integral in Eq. 10. 4 goes to 0 uniformly for all modes of approach of $\Gamma_{\epsilon_i,\eta_j}$ to Γ when $U=U''$.

We have verified that $U=U''$ is the solution in the sense F of $\delta J=0$. On taking $\delta U\equiv U$, as is permissible, it is clear that $L_{\epsilon_i,\eta_j\to 0}\delta J_{\epsilon_i,\eta_j}=2J$. Hence U'' yields an absolute, and not merely a relative, minimum for J. The variational problem and that of the homogeneous differential equation system, i. e. Eqs. 5, 7, 7. 1 and 7. 21, are one whence the conclusion

$$U' = U'' = C + sx + ty. \tag{10.7}$$

Since u_a, v_a, T_1, T_2, p, $\alpha = 1, 2$, involve U through its second derivatives only, their uniqueness is established.

We consider the case

$$(11) \qquad z = c\, S\frac{x}{a} S\frac{y}{a}.$$

Method I. The first method of determining a solution of the clamped sheet problem subject to Eq. 11 utilizes Eqs. 5, 6, 7, 7.1, 7.21. The demonstration of the validity of the ensuing formal developments is reserved to a later position in the paper.

Consider first Eq. 5. A possible solution is

$$(12) \qquad {}_1T = -\, d^2/8a^2[C\,2x/a + C\,2y/a]$$

where $d = C[A(1-\sigma^2)]^{1/2}$

$$(13) \qquad {}_1T|_\Gamma = -\, d^2/8a^2[2c^2\, y/a] \text{ or } -\, d^2/8a^2[2C^2\, x/a].$$

We wish next to determine ${}_2T$, a harmonic function, such that

$${}_1T|_\Gamma = -\, {}_2T|_\Gamma.$$

The formulae and later convergence discussion are simplified by replacing $C^2\dfrac{w}{a}$, $w = x, y$ of Eq. 13 by $1 - S^2\dfrac{w}{a}$

$${}_2T = d^2/8a^2\left[2 + 2/\pi \sum_{n=1,3\ldots}^{\infty} \frac{8}{m(m^2-4)} P \frac{Ch\dfrac{m}{a}\left(\dfrac{a\pi}{2} - y\right)}{Ch\dfrac{m\pi}{2}} S\frac{mx}{a}\right].$$

We write

$$(14) \quad T|_{x=0,a\pi} = (T - ({}_1T + {}_2T))|_{x=0,a\pi} = {}_3T|_{x=0,a\pi} = 2d^2/8a^2\pi \sum_m h_m S\frac{m}{a}y.$$

The symmetry of the problem requires summation over *odd* subscripts alone. Throughout this paper all summations are to be understood in this sense. For the sides $y = 0, a\pi$ it is merely necessary to interchange y and x in Eq. 14. The crux of our problem is essentially the determination of the values $\{h_m\}$.

We now exhibit T, i. e., ${}_1T + {}_2T + {}_3T$

$$(15) \qquad T = d^2/8a^2[2 - C\,2x/a - C\,2y/a + 2/\pi \Sigma P A_m f_m(x) S\frac{m}{a}y]$$

where $A_m = h_m + 8/m(m^2-4)$

$$f_m(w) = Ch\frac{m}{a}\left(\frac{a\pi}{2} - w\right)/Ch\frac{m\pi}{2}.$$

On making use of the Green's function, for instance, it is not difficult to find a solution of $\nabla^2 V = T$, namely

$$(16)\quad V = d^2/8a^2 \left[P \left\{ \frac{a^2}{4} C\, 2x/a + \frac{(x - a\pi/2)^2}{2} - 1/\pi \Sigma \frac{a}{m} A_m g_m(x)\, S \frac{m}{a} y \right\} \right]$$

where

$$g_m(w) = \frac{w\, Sh \frac{m}{a}(a\pi - w) + (a\pi - w)\, Sh \frac{m}{a} w}{2\, Ch^2 \frac{m}{2} \pi}.$$

Manifestly, $U = V + W$ also satisfies $\nabla^2 U = T$ provided W is a harmonic function. We determine this harmonic function by the conditions Eqs. 7 and 7. 1.

$$(7.01)\qquad (W_{xx} - \rho T + V_{xx})\big|_{x=0,a\pi} = 0$$

$$(7.11)\qquad (W_{xx} - (1-\rho)T + V_{xx})\big|_{y=0,a\pi} = 0.$$

Assuming the boundary values are taken on continuously, as indeed we show subsequently, we may write

$$(17)\qquad V_{xx}\big|_{x=0,a\pi} = \frac{d^2}{8a^2} \left[\frac{2}{\pi} \Sigma A_m\, S \frac{m}{a} y \right]$$

$$T\big|_{x=0,a\pi} = \frac{d^2}{8a^2} \left[\frac{2}{\pi} \Sigma h_m\, S \frac{m}{a} y \right] \quad \text{Cf. Eq. 14.}$$

Observing that

$$\frac{2}{\pi} \Sigma \frac{8}{m(m^2-4)} S \frac{m}{a} y = C \frac{2y}{a} - 1$$

we have therefore

$$(17.1)\qquad W_{xx}\big|_{x=0,a\pi} = \frac{d^2}{8a^2} \left[C \frac{2y}{a} - 1 + (1-\rho) \frac{2}{\pi} \Sigma h_m\, S \frac{m}{a} y \right].$$

Similarly

$$(17.2)\qquad W_{xx}\big|_{y=0,a\pi} = \frac{d^2}{8a^2} \left[C \frac{2x}{a} - 1 + (1-\rho) \frac{2}{\pi} \Sigma h_m\, S \frac{mx}{a} \right].$$

We are faced with the problem of determining a harmonic function subject to the condition that certain second derivatives take on prescribed values on the boundary. This seems somewhat different in form from the Dirichlet or Neumann problem of potential theory, but the solution is easily found in our simple case from the observation that W_{xx} must itself be harmonic. Explicitly then,

$$(18)\quad W_{xx} = \frac{d^2}{8a^2} \left[\frac{2}{\pi} \Sigma \left\{ (1-\rho) h_m + \frac{8}{m(m^2-4)} \right\} Q f_m(y)\, S \frac{mx}{a} \right] = - W_{yy}.$$

12

The value of W derived from Eq. 18 is subject to the arbitrary inclusion of

$$rxy + sx + ty + C.$$

Symmetry conditions require $r = s = t = 0$ leaving open the value of C as has already been remarked in the discussion of the uniqueness of U. (So far as the tensional stresses are concerned, the values of s, t, C are unimportant, and a non-vanishing r would make its presence evident in U_{xy} alone).

For purposes of completeness it is desirable to give the explicit expressions for T_1, T_2, S_1. Thus

$$T_1(x, y) = T_2(y, x)$$

$$(19)\quad T_1(x, y) = \frac{d^2}{8a^2}\Big[1 - C\frac{2y}{a} + \frac{2}{\pi}\Sigma\rho h_m f_m(y)\, S\frac{m}{ax}$$
$$+ (A_m - \rho h_m) f_m(x)\, S\frac{my}{a} + \frac{m}{2l} Q A_m g_m(x)\, S\frac{my}{a}\Big]$$

$$-S_1{}^{11} = U_{xy} = \int_{a\pi/2}^{x} U_{xxy}\,dx$$
$$= \frac{d^2}{8a^2}\Big[\Sigma\frac{2}{\pi}P\Big\{(A_m - 2\rho h_m)\tfrac{1}{2}\frac{Sh\frac{m}{a}\left(\frac{a\pi}{2} - y\right)C\frac{mx}{a}}{Ch\, m\pi/2}$$
$$+ A_m\tfrac{1}{2}\left(\frac{mx}{a}f_m(x)\, C\frac{my}{a} - m\pi\, Ch\frac{mx}{a}\, C\frac{my}{a}\right)\Big/2\, Ch^2\frac{m\pi}{2}\Big\}\Big]$$

Eq. 7.2 would yield series (and equations) in mh_m rather than in h_m. More seriously however, the series $mh_m\, S\frac{my}{a}\, q_m(x)$, $q_m = f_m$ or $\frac{m}{a}g_m$, requires conditions for its introduction which we are not prepared to justify. Accordingly, we replace Eq. 7.2 by Eq. 7.21

$$(7.21)^{12}\quad \int_{a\pi/2}^{y}\left(\frac{\partial T}{\partial x} + lU_{xxx}\right)\Big|_{x=0,a\pi} dy = 0$$
$$= \frac{d^2}{8a^2}\Big[\frac{2}{\pi}\Sigma(A_m + l\rho h_m)\, Th\frac{m\pi}{2}\, C\frac{my}{a}$$
$$+ \left(A_m + \frac{8l}{m(m^2-4)} + l(1-\rho)h_m\right)\int_{a\pi/2}^{y} f_m(y)\,dy\,\frac{m}{a}$$
$$+ \frac{l}{2}A_m\Big[Th\frac{m\pi}{2} + \frac{m\pi}{2\, Ch^2\, m\pi/2}\Big]C\frac{my}{a}$$
$$+ \frac{l}{2}A_m\int_{a\pi/2}^{y} g_m(y)\,dy\left(\frac{m}{a}\right)^2\Big].$$

---

[11] The symmetry in x and y of U_{xy} eliminates extraneous y functions introduced by the integration.

[12] The term by term differentiations here and earlier are justified by the uniform convergence of the resulting series which are included in Eq. 22.

In view of elementary properties of hyperbolic functions and the relation $l(1-\rho) = -(1+l)$ we may write this in the form

$$(19.1)\quad -\Big\{+\Sigma h_m\Big(1+\frac{l}{2}+l\rho+\frac{l}{2}\,\frac{m\pi}{Sh\,m\pi}\Big) + \Big(1+\frac{l}{2}+\frac{l}{2}\,\frac{m\pi}{Sh\,m\pi}\Big)\frac{8}{m(m^2-4)}\Big\}\,Th\,\frac{m\pi}{2}\,C\,\frac{my}{a}$$
$$=\Sigma\,\frac{8}{m^2-4}\,\frac{1+l}{a}\int_{a\pi/2}^{y} f_m(y)\,dy-\frac{lm}{a}\,h_m\int_{a\pi/2}^{y} f_m(y)\,dy + \frac{4lm}{a^2(m^2-4)}\int_{a\pi/2}^{y} g_m(y)\,dy+\frac{l}{2}\,\frac{m^2}{a^2}\,h_m\int_{a\pi/2}^{y} g_m(y)\,dy.$$

The Fourier series for f_m and g_m, assumed to be odd functions, and for their integrals are

$$f_m(y)=\sum_p\frac{4}{\pi}\,\frac{p}{(m^2+p^2)}\,S\,\frac{py}{a},\qquad \int_{a\pi/2}^{y} f_m(y)\,dy=-\Sigma\,\frac{4a}{\pi}\,\frac{1}{(m^2+p^2)}\,C\,\frac{py}{a}$$
$$g_m(y)=\sum_p\frac{8mpa}{\pi(p^2+m^2)^2}\,S\,\frac{py}{a},\qquad \int_{a\pi/2}^{y} g_m(y)\,dy=-\Sigma\,\frac{8}{\pi}\,\frac{ma^2}{(p^2+m^2)^2}\,C\,\frac{py}{a}.$$

On using these expansions we obtain on the right side of Eq. 19.1 certain double summations $\sum\limits_m\sum\limits_p$. We first interchange the order of summing. Then since the summation indices are "dummies" in the sense of tensor analysis, the scripts m, p may be replaced by p, m respectively. Accordingly, the right side of Eq. 19.1 becomes

$$(19.2)\quad \sum_m C\,\frac{my}{a}\sum_p\Big\{(1+l)\,\frac{4}{\pi}\Big\}\frac{8}{p^2-4}\,\frac{1}{p^2+m^2} + \frac{4l}{\pi}\,\frac{p^2}{(p^2+m^2)^2}\,\frac{1}{p^2-4}-\frac{4l}{\pi}\,h_p\,\frac{p}{p^2+m^2}+\frac{4l}{\pi}\,h_p\,\frac{p^3}{(p^2+m^2)^2}\Big\}.$$

The two sets of terms involving h_p may be combined to yield

$$-\frac{4l}{\pi}\sum_m\sum_p h_p\,\frac{m^2p}{(m^2+p^2)^2}.$$

To evaluate the summations involving $\dfrac{1}{p^2-4}$ we use contour integrals

$$(20)\quad \frac{1}{4i}\int_C\frac{1}{(z^2-4)}\,\frac{1}{(z^2+m^2)}\quad T\,\frac{z\pi}{2}\,dz=\sum_p\frac{1}{p^2-4}\,\frac{1}{p^2+m^2}\qquad (p=1,3,\cdots)$$

$$(20.1)\quad \frac{1}{4i}\int_C\frac{1}{z^2-4}\,\frac{z^2}{(z^2+m^2)^2}\quad T\,\frac{z\pi}{2}\,dz=\Sigma\,\frac{1}{p^2-4}\,\frac{p^2}{(p^2+m^2)^2}.$$

The integration path C may be taken as two Hankel loops. The first starts at $\infty + \epsilon i$, approaches its vertex at $z = \frac{1}{2}$ through valves in the first quadrant, then doubles back below the real axis to $\infty - \epsilon i$. The second is the reflection of this in the imaginary axis.[13] The loops may be closed by verticals away from the poles at $z = \pm mi$ and by allowing these gates to approach $+\infty$, $-\infty$ respectively the relations Eqs. 20 and 20.1 are easily verified. The proof that the path integrals remain finite is simple. We observe now that the integrals may also be evaluated by considering the residues associated with the singularities along the imaginary axis. When this is done, there is obtained

$$(21) \qquad \sum_p \frac{1}{p^2 - 4} \frac{1}{p^2 + m^2} = -\frac{\pi}{4m} \frac{1}{(m^2 + 4)} Th \frac{m\pi}{2}$$

$$(21.1) \qquad \sum_p \frac{1}{p^2 - 4} \frac{p^2}{(p^2 + m^2)} = \frac{\pi}{4m} Th \frac{m\pi}{2} \left\{ \frac{1}{2(m^2 + 4)} - \frac{4}{(m^2 + p^2)^2} \right\} - \frac{\pi^2}{16} Sch^2 \frac{m\pi}{2} \frac{1}{m^2 + 4}.$$

Eq. 19.1 is brought to the form

$$(19.3) \qquad \Sigma \gamma_m C \frac{m}{a} y = 0.$$

If the series on the left is a Fourier cosine series, it is well known that it must be unique. Hence the coefficients may be separately equated to 0. Thus

$$(19.4) \qquad \left[h_m \left(1 + l\rho + \frac{l}{2}\left(1 + \frac{m\pi}{Sh\, m\pi}\right)\right) + \frac{16m}{m^4 - 16}\left(1 + \frac{l}{2}\left(1 + \frac{m\pi}{Sh\, m\pi}\right)\right) + \frac{32l}{m(m^2 + 4)^2} + \frac{4l}{\pi} \frac{m^2}{Th \frac{m\pi}{2}} \sum_p \frac{p h_p}{(m^2 + p^2)^2} \right] = 0.$$

---

[13] Alternatively, the contours may be taken as the two lines $z = \pm \frac{1}{2}$. Another method of calculation, used as a check, is based on an extension of the usual Parseval theorem. For the applications we require the algorithm, certainly valid for functions of class C',

$$\int_{-\pi}^{\pi} F_1(t) F_2(x - t)\, dt = \pi \Sigma a_m C m x = \pi F(x)$$

where

$$a_m = b_m c_m, \qquad F_1(x) = \Sigma b_m \frac{Smx}{Cmx}, \qquad F_2(x) = \Sigma c_m \frac{Smx}{Cmx},$$

The procedure is summed up in the formula

$$2/\pi \int_{-\pi}^{\pi} F(x) G(x)\, dx = \Sigma a_m g_m, \qquad G(x) = \Sigma g_m Cmx.$$

It enables us to sum series such as

$$\Sigma_m [(m^2 \pm a^2)(m^2 \pm p^2)(m^2 \pm r^2)]^{-1}, \qquad a,\ p,\ r \text{ real}$$

for the functions, associated with $(m \pm a^2)^{-1}$ etc. as Fourier coefficients, are known.

On replacing l, ρ by the σ equivalents, we have

$$(19.5)\quad \left[h_m\left(3-\sigma-(1+\sigma)\frac{m\pi}{Sh\, m\pi}\right)-\frac{8}{\pi}\frac{(1+\sigma)}{Th\, m\pi/2}m^2\sum_p\frac{ph_p}{(m^2+p^2)^2}\right.$$
$$\left.+\left[(3-\sigma)-(1+\sigma)\frac{m\pi}{Sh\, m\pi}\right]\frac{16m}{m^4-16}-\frac{64(1+\sigma)}{m(m^2+4)^2}=0.\right.$$

This is the central equation.

The legitimacy of the procedure of the Method I will now be investigated. We must justify (a) obtaining the boundary values of the functions defined by series by introducing $x=0, a\pi$ $(y=0, a\pi)$ in the terms of the series,[14] (b) the grouping of terms of the several series, (c) the interchange of summation order in m and p, and (d) the assertion that the sum on the left side of Eq. 19. 3 is a Fourier series.

We anticipate the key inequality of Eq. 38, namely

$$(38)\qquad |h_m| \leq E/m^{1+\delta},\quad 1>\delta>0 \quad \text{for} \quad \sigma<5/7.3$$

and it is for such σ values, alone, that the following demonstrations apply. The series we have to consider are of the form

$$(22)\qquad a_m q_m(w)p_m(t) \qquad w, t = x, y \text{ or } y, x$$

$q_m(w)$ represents alternatively $f_m(w)$, $\frac{m}{a}g_m(w)$, $\frac{a}{m}f'_m(w)$, $g'_m(w)$,

$$\int_{a\pi/2}^{w}\frac{m}{a}f_m(w)\,dw, \int_{a\pi/2}^{w}\left(\frac{m}{a}\right)^2 g_m(w)\,dw.$$

The prime denotes differentiation. $p_m(t)$ is either $S\left(\frac{m}{a}t\right)$ or $C\left(\frac{m}{a}t\right)$. a_m stands for h_m or $1/m(m^2-4)$, $1/m(m^2+4)$, $1/m(m^2+4).^2$ We remark two important characteristics: (i) $(q_m(w)p_m(t))$ is of class C^0 in the closed square. (ii) $|q_m(w)p_m(t)| < M$, $0 \leq {w \atop t} \leq a\pi$.

Evidently $\Sigma\,|a_m|$ is a convergent series according to Eq. 38. It then follows from (ii) that the series of type Eq. 22 converge absolutely and uniformly in the closed square, thus justifying operations (a) and (b).

Furthermore, in view of (i) the functions defined by the series are of class C^0. Starting with Eqs. 19 we may easily write down the explicit formal

[14] It is elementary of course that the Fourier series of a function with a discontinuity at $x=0$ does not approach the value of the series at the point from either side. Thus, for instance, Eq. 14 does not really posit 0 values for T at the corners. This fact is rather a product of the analysis and is connected with the uniform convergence of the series Eq. 22.

expansion of the right hand integral, exclusive of limit sign, in Eq. 7.3. The expansion in question is a sum of series of type Eq. 22 where $q_m(w)$ is restricted to the last four expressions catalogued above. It is accordingly of class C^0 as was asserted in connection with convention F.

We now consider (c). The series

$$\Sigma\Sigma a_m \frac{m}{m^2+p^2} C\frac{p}{a}y, \qquad \Sigma\Sigma a_m \frac{m^3}{(m^2+p^2)^2} C\frac{py}{a}$$

are evidently absolutely convergent double series. Therefore the order of summation may be interchanged, as was done to obtain Eq. 19.2.

We proceed to (d). Each term in the left hand bracket of Eq. 19.4 is dominated by $M/m^{1+\delta}$. The only term for which this assertion is not immediately evident is

$$R_m = 1/Th\, m\pi/2 \sum_p m^2 p h_p/(m^2+p^2)^2.$$

However, on writing $|h_m| \leqq E/m^{1+\delta}$ in place of Eq. 35, it is clear that Eqs. 36, 36.1 and 37, cf. sequel, establish this sort of dominant for

$$R_m\left(1 - \frac{1+\sigma}{3+\sigma}\,\frac{m\pi}{Sh\, m\pi}\right)$$

and accordingly for R_m as well. Hence γ_m, i.e., the left hand bracket of Eq. 19.4 is inferior to $\gamma/m^{1+\delta}$ where γ is a suitable fixed constant. Therefore

$$\Sigma\, \gamma_m{}^2 < \infty.$$

The Riesz Fisher theorem now guarantees that the left hand side of Eq. 19.3 is a Fourier series.

Our derivation of Eq. 19.4 is therefore unexceptionable.

Method II. Eq. 19.5 (19.4) may also be arrived at by a different method, which besides its independent interest incidentally affords a check on the accuracy of that equation. Except for possible occurrence of sines instead of cosines the series encountered below are of the types Eqs. 22.1—22.3. Accordingly, the validation of the operations conducted in the development of Method II is really already implicitly subsumed in the discussion of those equations.

The starting point now is Eq. 3.1 which may be written in the form

$$\nabla^2 u = -\frac{1}{A(1-\sigma)}\,\frac{\partial T}{\partial x} - z_x \nabla^2 z. \tag{3.11}$$

If the last term on the right be dropped, a solution of the resulting equation is evidently

$$(23)\quad u_1 = -\frac{1}{A(1-\sigma)}\frac{\partial V}{\partial x}$$
$$= (1+\sigma)c^2/8a^2\left[\frac{a}{2}\left(S\frac{2x}{a} - \frac{x - a\pi/2}{a}\right)\right.$$
$$\left. + \frac{1}{\pi}\Sigma A_m \frac{a}{m}\left[g_m(y)\, C\frac{mx}{a}\frac{m}{a} + S\frac{m}{a}\, y g'_m(x)\right]\right] \text{ cf. Eqs. 11, 17.}$$

A solution of

$$\nabla^2 u = -z_x \nabla^2 z = \frac{c^2}{2a^3}\left[S\frac{2x}{a} - S\frac{2x}{a}C\frac{2y}{a}\right]$$

is

$$(23.1)\qquad u_2 = -\frac{c^2}{8a^2}\left[a\left\{S\frac{2x}{a} - \tfrac{1}{2}S\frac{2x}{a}C\frac{2y}{a}\right\}\right].$$

We require now a harmonic function u vanishing for $y = 0, a\pi$ and such that

$$u_1 + u_2|_{x=0} = -(u_1 + u_2)|_{x=a\pi} = -u_3|_{x=0} = u_3|_{x=a\pi}.$$

On making use of Eq. 20 we have

$$(24)\quad u_3 = -\frac{c^2}{8a^2}(1+\sigma)\left[\left(\frac{a\pi}{2} - x\right) + \frac{1}{\pi}\Sigma\Sigma\frac{a}{m}A_m\left\{\frac{8m^2p}{\pi(m^2+p^2)^2}\right.\right.$$
$$\left.\left.\times\frac{Sh\frac{p}{a}\left(\frac{a\pi}{2} - x\right)}{Ch\frac{p\pi}{2}}S\frac{py}{a} + \frac{Sh\, m\pi + m\pi}{2Ch^2\frac{m\pi}{2}}\,\frac{Sh\frac{m}{a}\left(\frac{a\pi}{2} - x\right)}{Ch\frac{m\pi}{2}}S\frac{my}{a}\right\}\right]$$

$u_1 + u_2 + u_3$ vanishes on $x \to 0, a\pi$ and takes on the value

$$(24.1)\qquad c^2/8a^2\left[\frac{a}{2}\sigma S\frac{2x}{a}\right]$$

for $y = 0$, $a\pi$. To balance this we introduce the harmonic function u_4 which takes on the negative of the values in Eq. 27.1 for $y \to 0$, $a\pi$ and is 0 on the other two sides, viz.

$$(25)\qquad u_4 = -\frac{c^2}{8a^2}\left[\frac{a\sigma}{2}S\frac{2x}{a}Ch\frac{2}{a}\left(\frac{a\pi}{2} - y\right)/Ch\,\pi\right]$$

The required solution of Eq. 3.11 with $u|_\Gamma = 0$ is

$$(26)\quad u(x,y) = u_1 + u_2 + u_3 + u_4 = v(y,x)$$

$$= \frac{c^2}{8a^2}\Bigg[(\sigma - 1)\, C\,\frac{2x}{a} + C\,\frac{2x}{a}\, C\,\frac{2y}{a} - \sigma\, C\,\frac{2x}{a} \left\{ \frac{Ch\,\frac{2}{a}\left(\frac{a\pi}{2} - y\right)}{Ch\,\pi} \right\}$$

$$+ \frac{1+\sigma}{\pi} \sum_m \sum_p \frac{a}{m} A_m \Bigg\{ g_m(y)\, C\,\frac{mx}{a}\,\frac{m}{a} + S\,\frac{my}{a}$$

$$- \frac{\left[Sh\,\frac{m}{a}\,(a\pi - x) - Sh\,\frac{m}{a}\,x - \frac{mx}{a}\, Ch\,\frac{m}{a}\,(a\pi - x) + (a\pi - x)\,\frac{m}{a}\, Ch\,\frac{m}{a} \right.}{2Ch^2\,\frac{m\pi}{2}}$$

$$- \frac{8m^2p}{\pi(p^2+m^2)^2} \left(\frac{Sh\,\frac{p}{a}\left(\frac{a\pi}{2} - x\right)}{Sh\,\frac{p\pi}{2}} \right) S\,\frac{py}{a}$$

$$- \frac{Sh\,m\pi + m\pi}{2\,Ch^2\,\frac{m\pi}{2}}\; \frac{Sh\,\frac{m}{a}\left(\frac{a\pi}{2} - x\right)}{Sh\,\frac{m\pi}{2}}\, S\,\frac{my}{a} \Bigg\} \Bigg].$$

We recall that

$$u_x + z_x{}^2/2\,|_{x=0,a\pi} = \frac{1}{A(1+\sigma)}\, T\,|_{x=0,a\pi}$$

$$= (1-\sigma)\,\frac{c^2}{8a^2}\, \Sigma h_m\, S\,\frac{my}{a}.$$

Thus after some direct reduction

$$(27)\quad \Bigg[2\,S^2\,\frac{y}{a} + \sigma\left(1 - \frac{Ch\,\frac{2}{a}\left(\frac{a\pi}{2} - y\right)}{Ch\,\pi}\right) + \frac{1+\sigma}{\pi} \sum_m \sum_p A_m$$

$$\times \left\{ -2\,S\,\frac{my}{a} + \frac{8mp^2}{\pi(m^2+p^2)^2}\,\frac{1}{Th\,\frac{p\pi}{2}}\; \frac{S\,\frac{py}{a} + Sh\,m\pi - m\pi}{2\,Ch^2\,\frac{m\pi}{2}}\; \frac{1}{Th\,\frac{m\pi}{2}}\, S\,\frac{my}{a} \right\} \Bigg]$$

$$= (1-\sigma)\,\frac{2}{\pi} \sum h_m\, S\,\frac{my}{a}.$$

The regrouping of terms taken from the set $\{u_r\}$, $r = 1 \cdots 4$ is allowable here since all the series entering are absolutely convergent in view of Eq. 38. The summation order on p and m may be interchanged and the scripts p, m substituted for m, p by the absolute convergence property of

$$\Sigma\,\Sigma\, A_m m p^2 \Big/ \left[(m^2+p^2)^2\, Th\,\frac{m\pi}{2} \right] S\,\frac{py}{a}.$$

By using a contour integral of the type introduced in Eqs. 20, 20.1 or the extended Parseval theorem, we may show that

$$(20.2)\quad m^2\sum_p \frac{1}{p^2-4}\,\frac{1}{(p^2+m^2)^2} = -\frac{\pi}{8m}\left\{Th\,\frac{m\pi}{2}\left[\frac{3m^2+4}{(m^2+4)^2}\right]\right\}+\frac{\pi^2}{16}\,Sch^2\,\frac{m\pi}{2}\,\frac{1}{(m^2+4)}.$$

On replacing the terms outside the summation sign in Eq. 27 by their absolutely and uniformly convergent Fourier sine expansions, we may express that equation as

$$(27.1)\qquad \sum \beta_m\, S\,\frac{m}{a}\,y=0.$$

Absolute convergence of the sine series whose coefficients are

$$1/m(m^2+4),\quad 1/m(m^2-4),\quad m^2\sum_p h_p p/(m^2+p^2)^2,\quad h_m$$

justifies the attendant groupings. Furthermore, since by Eq. 38 the coefficients in each of these series go down at least as rapidly as $m^{-1-\delta}$ we conclude that

$$|\beta_m|\leq \beta/m^{1+\delta}\quad \text{for}\quad \sigma<5/7.3,\quad 1>\delta>0.$$

In the same manner as in the case of Eq. 19.3, we show that the left side of Eq. 27.1 is a Fourier *sine* series and hence that $\beta_m\equiv 0$ where $\beta_m(\equiv\gamma_m)$ is precisely the expression in Eq. 19.5.

In view of our use of convention F and in the interests of completeness, we consider the displacements on the basis of the stresses derived in Method I. Incidentally we are enabled to exhibit an alternative expression for $u(x,y)$ somewhat simpler in form than that of Eq. 26. In fact Eq. 1 yields

$$\int_{a\pi/2}^{x}(u_x+z_x^2/2)\,dx=(A(1-\sigma^2))^{-1}\int_{a\pi/2}^{x}(T_1-\sigma T_2)\,dx$$

$$(28)\quad u(x,y)=c^2/8a^2\Bigg[\sigma\left(\frac{a\pi}{2}-x\right)+\frac{a\sigma}{2}\,S\,\frac{2x}{a}-a\,S\,\frac{2x}{a}\,S^2\,\frac{y}{a} + \frac{2}{\pi}\sum\Bigg\{\sigma\alpha_m\,\frac{a}{m}\,f_m(y)\,C\,\frac{mx}{a}+\left(\alpha_m\left(\frac{3+\sigma}{2}\right)+h_m\left(\frac{3-\sigma}{2}\right)\right)f'_m(x)\,S\,\frac{my}{a} + \frac{m}{a}\,\frac{1+\sigma}{2}\,A_m\left[\left(-xf_m(x)+\frac{a\pi\,Ch\,\frac{mx}{a}}{2\,Ch^2\,\frac{m\pi}{2}}\right)S\,\frac{my}{a}+g_m(y)\,C\,\frac{mx}{a}\right]\Bigg\}\Bigg]$$

where $\alpha_m=8/m(m^2-4)$. The earlier reasoning, involving Eq. 38, establishes that $u(x,y)$, $u_x(x,y)$ and $u_y(x,y)$ are represented by absolutely and

uniformly convergent series of continuous functions and are therefore continuous in the closed square. Moreover the termwise integrations and differentiations leading to $u_y(x, y)$ are valid.

Now, $u_y(x, y)|_{x=0,a\pi}$ turns out to be identical with the right hand side of Eq. 7. 21 and therefore vanishes. That is to say, $u(x, y)$ is constant on $x = 0, a\pi$. The continuity of $u(x, y)$ justifies our writing

$$(29) \qquad u(0, y) = u(0, 0) = c^2/8a^2[\sigma a\pi/2 + 2/\pi\Sigma\sigma \frac{a}{m} \alpha_m].$$

Furthermore,

$$(29.1) \quad u(x,0) = c^2/8a^2[\sigma(a\pi/2 - x) + \sigma a/2\, S\, 2x/a + 2/\pi\Sigma\, (\sigma\alpha_m a/m) C\, mx/a].$$

The summations in each of Eqs. 29 and 29. 1 are easily evaluated, by integrating the sine series for $S^2\, x/a$, whence it is seen that the right hand sides of these equations vanish. Similar deductions apply for the sides $x, y = a\pi$. We have thus established that

$$u(x, y)|_\Gamma = v(y, x)|_\Gamma = 0.$$

Since $u_y(x, y)$, $v_x(x, y)$ and $z_x z_y$ vanish at the corners, evidently S_1 (cf. Eqs. 1. 1) vanishes at the corners. From Eqs. 19, however, we should get for the corner 0, 0

$$-S_1(0, 0) = 2d^2/8a^2 P\Sigma\{(A_m - 2ph_m) Th\, m\pi/2 - A_m \frac{m\pi}{4} Sch^2\, m\pi/2\}.$$

The prescribed vanishing of the right hand expression above furnishes a check on the computed values of h_m.

Solution of the algebraic system. We turn now to the consideration of the solution of Eq. 19. 5, which may be written for convenience in the form [15]

$$(31) \qquad h_m - K_m(\sigma)\Sigma\alpha_{mr}h_r = C_m(\sigma)$$

where

$$K_m(\sigma) = b(\sigma)/Th \frac{m\pi}{2}\left(1 - G(\sigma)\frac{m\pi}{Sh\, m\pi}\right), \quad \alpha_{mr} = m^2 r/(m^2 + r^2)^2$$

$$C_m(\sigma) = \frac{H_m}{m(m^2+4)^2} - \frac{16m}{m^4 - 16}\left[1 + \frac{g}{1 - Gm\pi/Sh\, m\pi}\right]$$

$$b(\sigma) = \frac{8}{\pi}\frac{1+\sigma}{3-\sigma}, \quad G(\sigma) = \frac{1+\sigma}{3+\sigma}, \quad g = \frac{2\sigma}{3-\sigma}, \quad H_m = 8b(\sigma)\pi/\left(1 - G\frac{m\pi}{Sh\, m\pi}\right)$$

We wish to bound $K_m(\sigma)\Sigma\alpha_{mr}$. Clearly,

[15] The positiveness of α_{mr}, $K_m(\sigma)$, β_{mr} allows us to dispense with absolute value signs, but is in no wise essential for the proofs.

(32) $$m^2\Sigma r/(m^2+r^2)^2 \leq \Sigma r/(1+r^2)^2.$$

Now

$$\Sigma r/(1+r^2)^2 - \frac{1}{4} - \frac{3}{100} - \frac{5}{(26)^2} - \frac{7}{(50)^2} \leq \tfrac{1}{2}\int_7^\infty \frac{rdr}{(1+r^2)^2} = .005.$$

Therefore the left side of Eq. 32 is inferior to .282. We note

$$K_m(\sigma) \leq K_1(\sigma).$$

The condition

(32. 1) $$K_m(\sigma) \sum_r \alpha_{mr} \leq w < 1$$

is satisfied for $w = K_1(\sigma)(.282) < 1$. Since $K_1(\sigma)$ is an increasing function of σ from 0 to 1 and $K_1(19/20)(.282) < 1$, the inequality Eq. 32. 1 is verified for $\sigma \leq 19/20$.

The restriction of interest at this stage is $|h_m| \leq M$. Evidently $|C_m(\sigma)|$ is inferior to a constant D for all n. The solution of Eq. 31 is unique if Eq. 32. 1 is satisfied, and may be represented by a Neumann expansion

(33) $$h_m = C_m + \sum_r K_m\alpha_{mr}C_r + \sum_r\sum_s K_m\alpha_{mr}K_r\alpha_{rs}C_s + \cdots$$

where the argument σ has been omitted in the C and K terms.[16]

One method of approximation, in principle, to the true value of h_m is that of using a finite number of terms in the expansion to the right of Eq. 33. The error committed in stopping at the n-th term is inferior to $Dw^n/1-w$, where $|C_m| < D < \infty$. Actually, however, although $K_m \sum_r \alpha_{nr}C_r$ may be calculated accurately, by using contour integration for instance, the value of $\sum_r\sum_s K_m\alpha_{nr}K_r\alpha_{rs}C_s$ is not so amenable.[17] Eq. 32. 1 guarantees the applicability of the practical method of finite segments.[18] This procedure amounts to setting $h_m = 0$ for $m \geq 2N+1$ and solving for the first N h's from the first N equations of Eq. 31 by Cramer's rule. The approximation obtained in this way denoted by h_m^N.

We make the application of the method of finite segments more precise by exhibiting bounds for the error in h_m^N. The solution of the N equations in N unknowns may, of course, also be written as a Neumann expansion when Eq. 32. 1 holds. We denote by k_m^N the solution of the set of equations obtained by replacing C_m by its absolute value. Accordingly

---

[16] March, *loc. cit.*, p. 311.

[17] The difficulty arises from the fact that the summations yield the sum of two logarithmic first derivatives of the gamma function, with conjugate complex arguments, which are not tabulated so far as the writer is aware. (The terms in $m\pi/Shm\pi$ give no trouble since they are negligible for large m).

[18] Pellet, *Bull. de la Soc. Math. de France*, t. 42 (1914), p. 48.

$$| k_m{}^N - k_m{}^n | = \sum_{2n-1}^{2N-1} \beta_{mr} | C_r | + \left(\sum_1^{2n-1} \sum_{2n-1}^{2N-1} \beta_{ms}\beta_{sr} | C_r | + \sum_{2n-1}^{2N-1} \sum_1^{2N-1} \beta_{ms}\beta_{sr} | C_r | \right) + \cdots$$

$$\geq \left| \sum_{2n-1}^{2N-1} \beta_{mr} C_r + \left(\sum_1^{2n-1} \sum_{2n-1}^{2N-1} \beta_{ms}\beta_{sr} C_r + \sum_{2n-1}^{2N-1} \sum_1^{2N-1} \beta_{ms}\beta_{sr} C_r \right) + \cdots \right.$$

(34) $$= | h_m{}^N - h_m{}^n | \text{ for } n < N$$

where

$$\beta_{mr} = K_m \alpha_{mr}.$$

Evidently

$$0 \leq k_m{}^N < D/1 - w.$$

Thus for m fixed $\{k_m{}^N\}$ is a uniformly bounded non-decreasing sequence, and has a unique limit value k_m. Therefore

(34.1) $$| k_m - k_m{}^N | \geq | h_m - h_m{}^N |.$$

Furthermore

(34.2) $$| k_m | \geq k_m{}^N \geq | h_m{}^N |$$
$$| k_m | \geq | h_m |.$$

Clearly,

(35) $$| C_m | < D/m^{1+\delta}$$

where $1 > \delta > 0$ and the minimum value for D depends on σ. We have then

(36) $$\sum_p K_m \alpha_{mp}/p^{1+\delta} < K_1(\sigma) m^2 \sum 1/p^\delta (m^2 + p^2)^2$$
$$< \frac{K_1(\sigma) m^2}{2} \int_0^\infty \frac{dp}{p^\delta (m^2 + p^2)^2}$$

for

$$\tfrac{1}{2} \int_0^1 \frac{dp}{p^\delta (m^2 + p^2)^2} > \tfrac{1}{2} \int_0^1 \frac{dp}{m^2 + p^2} > \frac{1}{(m^2 + 1)^2}.$$

Now

(36.1) $$m^2 \int_0^\infty \frac{dp}{p^\delta (m^2 + p^2)^2} = \frac{1}{2m^{1+\delta}} \frac{\Gamma\left(\frac{3+\delta}{2}\right) \Gamma\left(\frac{1-\delta}{2}\right)}{\Gamma(2)}.$$

We have

$$\Gamma\left(\frac{3+\delta}{2}\right) = \frac{1+\delta}{2} \Gamma\left(\frac{1+\delta}{2}\right)$$

$$\Gamma\left(\frac{1+\delta}{2}\right) \Gamma\left(\frac{1-\delta}{2}\right) = \pi/\sin \frac{1-\delta}{2} \pi = \pi/\cos \frac{\delta\pi}{2}.$$

Hence the right side of Eq. 33 is inferior to

(37) $$\frac{D}{m^{1+\delta}} \left[1 - K_1(\sigma) \frac{\pi}{8} \left(\frac{1+\delta}{\cos \frac{\delta\pi}{2}} \right) \right]^{-1} = E/m^{1+\delta}.$$

The foregoing argument involves

$$(37.1) \qquad 1 > K_1(\sigma)\,\frac{\pi}{8}\,\frac{1+\delta}{\cos\frac{\delta\pi}{2}} > K_1(\sigma)\,\frac{\pi}{8}.$$

This is satisfied by $1 > \delta > 0$ for $\sigma < 5/7.3$. Hence from Eqs. 34.2, 35 and 37.2

$$k_m < E/m^{1+\delta}$$

$$(38)^{19} \qquad |\,h_m\,|,\ |\,h_m{}^N\,|,\ k_m{}^N < E/m^{1+\delta},\quad 1 > \delta > 0,\ \sigma < .685.$$

For $\sigma = \frac{1}{4}$ a rough calculation yields $\delta = .272$.

For utility in actual computations we replace k_m, $k_m{}^N$ by quantities which can be calculated conveniently. We introduce the following system of dominant equations which can be solved in closed form

$$(39) \qquad l_m - b' \sum l_r \frac{r}{(r^2+m^2)\ Th\ m\pi/2} = D'/m.$$

The dominant property holds provided D' be taken large enough and

$$b' \sum_r (m^2+r^2)^{-1} \geq b m^2 \sum_r (m^2+r^2)^{-2}$$

for then the Neumann expansion for l_m is surely greater term by term than that for k_m. Since

$$\Sigma\,(m^2+r^2)^{-1} = \frac{\pi}{4m}\ Th\frac{m\pi}{2},\quad m^2\,\Sigma\,(m^2+r^2)^{-2} = \left[\frac{\pi}{8m} - \frac{\pi^2}{8\,Sh\ m\pi}\right] Th\frac{m\pi}{2}$$

we require

$$\frac{b'\pi}{4m} \geq b\left(\frac{\pi}{8m} - \frac{\pi^2}{8\,Sh\ m\pi}\right)$$

which is satisfied by

$$b' = b/2.$$

With this value for b the solution for l_m is

$$(40) \qquad l_m = \frac{D'}{m} \Big/ \left(1 - \frac{b\pi}{8}\right) \geq k_m$$

which, according to Eq. 38, is valid for $\sigma < .685$. On combining Eqs. 34.1 and 40 convenient majorants for the errors $(h_m - h_m{}^N)$, in the N-th approximation, are available, viz.:

---

[19] Since $k_m > 0$ and

$$L_{m\to\infty} \int_1^\infty \frac{dp}{p\,(p^2+m^2)^2} = O\,(\log m/m^4)$$

it may be shown that $\delta < 1$ in Eq. 38 even if Eq. 35 is replaced by $|\,C_m\,| < D/m^3$. This refers to the bound for k_m. However, there is a remote possibility that on taking account of the relation $C_1/C_m < 0$, $m > 1$ that $\delta > 1$ is available for bounding $|\,h_m\,|$ for some σ. The point of this observation lies in the then consequent direct availability of Eq. 7.2 instead of Eq. 7.21 in the sense F.

(34. 3) $$| D'/m \left(1 - \frac{b\pi}{8}\right) - l_m{}^n | \geq | h_m - h_m{}^N |, \quad n \leq N.$$

Numerical applications. For many materials $\sigma = \frac{1}{4}$. As illustration of the paragraphs immediately preceding as well as the technological interest of this case we summarize some rough calculations (for $\sigma = \frac{1}{4}$). The fourth stage finite segment equations are obtained from Eq. 19. 4;

$$\begin{aligned}
.34195h_1{}^4 - .02323h_3{}^4 - .005728h_5{}^4 - .0021683h_7{}^4 &= +\ 1.20707\\
-\ .06367h_1{}^4 + .55175h_3{}^4 - .02754h_5{}^4 - .01325h_7{}^4 &= -\ .498\\
-\ .02616h_1{}^4 - .03589h_3{}^4 + .57575h_5{}^4 - .0226052h_7{}^4 &= -\ .09065\\
-\ .01386h_1{}^4 - .03091h_3{}^4 - .03165h_5{}^4 + .58586h_7{}^4 &= -\ .03301.
\end{aligned}$$

The values for $h_1{}^4$, $h_3{}^4$ and $h_5{}^4$ are about

$$h_1{}^4 = 4.058, \qquad h_3{}^4 = -.434, \qquad h_5{}^4 = -.0075.$$

Manifestly, the best bounds of accuracy for $h_m{}^{(4)}$, consequent on Eq. 34. 1 are obtained for $n = N = 4$. For the purposes of this paper, however, we may content ourselves with $n = 2$ (and the slide rule). It is easy to show that for $\sigma = \frac{1}{4}$, D' may be taken as 2.54. From Eq. 39 we get

(42) $$\begin{aligned}
.587l_1{}^2 - .198l_3{}^2 &= 2.54\\
-\ .066l_1{}^2 + .89l_3{}^2 &= .925\\
l_1{}^2 = 4.7, \qquad l_3{}^2 &= 1.3.
\end{aligned}$$

Eq. 40 yields

(42. 1) $$l_1 = 5.75, \qquad l_3 = 1.92.$$

Hence the error in $h_1{}^2$, for the *two* stage calculation ~ 1 and that of $h_3{}^1 \sim .6$. The values obtained in Eq. 41. 1 involving a *four* stage approximation are better. Since $l_m = \dfrac{5.75}{m} > | h_m |$ it is easy to see from the expressions for u, v, T_1, T_2, and S_1 that h_1 and h_3 alone need be retained for *most* technological applications. This simplification enables us also to conveniently determine the general features of the pressure variations by substituting the reduced expressions for T and U_{xy} in Eq. 6.

We proceed to a derivation of the c, p relation consistent with the theory presented in this paper. This relation is the whole purpose of Foppl's analysis and he obtains

(43) $$c = .802(a_F)^{4/3}(p/E_F h)^{1/3} \quad \text{Cf. Eq. 11}$$

where $2a_F$ is the side of the square.

$\overline{p}$ replaces Foppl's $\dfrac{p}{h}$ for our treatment and is the mean pressure averaged with respect to vertical displacement. That is to say,

$$(43.1) \qquad \text{Work} = \int_0^{a\pi}\int_0^{a\pi} pz\,dxdy = \bar{p}\int_0^{a\pi}\int_0^{a\pi} z\,dxdy.$$

Combining Eqs. 6 and 43. 1 there results, after straightforward evaluation of integrals and the choice $\sigma = \frac{1}{4}$

$$(44) \quad -32a^4\pi\bar{p}/cd^2 = \Bigg(-\frac{3\pi^3}{4} + 8\Sigma\left[\frac{8}{m(m^2-4)} + 6h_m\right]\frac{2}{m^4-16}\,Th\,\frac{m\pi}{2}$$

$$+ A_m\Bigg\{\frac{8}{m^2(m^4-16)}\,Th\,\frac{m\pi}{2} - \frac{m\pi}{m^4-16}$$

$$+ \frac{4m^2}{(m^2+4)(m^4-16)}\,Th\,\frac{m\pi}{2} - m\pi\frac{Th^2\,\frac{m\pi}{2}}{m^4-16}\Bigg\}\Bigg).$$

A rough calculation indicates that the predominant contribution in the main bracket is that of $-3\pi^3/4$ and that the terms under the summation sum to about -10. Since cube roots are taken below, the contribution of these summation terms has very little influence. That is to say the values assigned h_m may be varied considerably without appreciably affecting the numerical coefficient in Eq. 44. 2 below. We have

$$(44.1) \qquad cd^2 = \left(\frac{3\pi^3}{4} + 10\right)/32\pi a^4\bar{p} = c^3A(1-\sigma^2)$$

$$(44.2) \qquad c = 1.4a^{4/3}\bar{p}^{1/3}.$$

To compare this with Foppl's value we note the relations

$$\bar{p} = (p/h)_F, \qquad a_F = a\pi/2, \qquad E_F = A(1-\sigma^2).$$

Since $\left(\frac{2}{\pi}\right)^{4/3} 1.4 \sim .8 +$ we have

$$(44.3) \qquad C \sim .8a_F^{4/3}(p/E_Fh)^{1/3}.$$

There is almost perfect concordance between Foppl's result and our own. In view of the insensitiveness to the values of h_m, this, in itself, has comparatively little significance in justifying his approximations. Cf. Eq. 28.

UNIVERSITY OF ILLINOIS,
URBANA, ILLINOIS.

TUBES AND SPHERES IN n-SPACES, AND A CLASS OF STATISTICAL PROBLEMS.*[1]

By Harold Hotelling.

1. The geometrical and the statistical problems. With reference to a curve C with continuously turning tangent in a metrical space of any number of dimensions, we define a *tube* as the locus of points at a fixed distance θ, called the *radius,* from C, the distance being measured in each case along a geodesic perpendicular to C. A *sphere* or *geodesic sphere* is of course the locus of points at a fixed geodesic distance from a given point. Lengths and areas of geodesic circles on a surface have been investigated by Bertrand and Diguet,[2] who obtained the first two non-vanishing terms in the expansion in powers of the radius. We shall generalize this result for spheres in n dimensions in § 5, and for tubes in § 6. The first term in such an expansion is independent of the curvature properties of the space and of the curve, and may therefore be found from the case of euclidean space. We shall see that alternate terms in the series vanish. The problem is then to express the others in terms of known invariants; this will be done for the first non-vanishing terms following the euclidean ones. We shall also find exact and simple expressions for the volumes enclosed by tubes in euclidean and spherical spaces. In both these cases the volume enclosed is exactly the product of the length of the curve by the $(n-1)$-dimensional area of a cross-section. The qualification must however be made that overlapping regions must be counted with their appropriate multiplicities. A necessary condition for non-overlapping will be obtained. We shall confine our consideration to spaces of positive definite distance elements.

A special type of normal coördinates associated with the arbitrarily given curve is introduced in § 6. These may prove useful in a variety of geometrical and physical problems.

Tubes on a hypersphere play a part in theoretical statistics. For example, if a set of observations

$$x_1, x_2, \cdots, x_n$$
$$y_1, y_2, \cdots, y_n$$

* Received August 22, 1938.

[1] Presented to the American Mathematical Society, September 6, 1938.

[2] *Journal de Mathématiques* (Liouville), Ser. 1, vol. 13 (1848), pp. 80-86. L. P. Eisenhart, Differential Geometry (1909), p. 209.

is used to determine the parameters b and p in the regression equation

$$Y = bf(x, p) \tag{1.1}$$

in such a way that

$$\sum_{a=1}^{n} \{bf(x_a, p) - y\}^2$$

is a minimum, then the correlation R between the fitted values and the observed values y (calculated without elimination of the mean) is the cosine of the angle made by two lines through the origin of cartesian coördinates in euclidean space of n-dimensions, one line through the point of coördinates y_a, the other through the point of coördinates Y_a $(a = 1, 2, \cdots, n)$. On the hypothesis that the y_a have no real relation to the x_a, but are normally and independently distributed about zero with a common variance, we may regard the y-line as drawn to a random point on a unit hypersphere, on which there is a uniform distribution of such points. The other line is drawn to a point on the curve whose equations in terms of the parameter p may be taken as

$$Y_a = f(x_a, p), \qquad (a = 1, \cdots, n).$$

or the functions on the right may be multiplied by any constant. The method of least squares is such that R is made a maximum; consequently the Y-line is drawn through a point of the unit hypersphere lying on the curve C into which the foregoing curve is projected from the center, in such a way that the geodesic distance between the intersections of the two lines with the hypersphere is a minimum. The probability that R exceeds any assigned or observed value is the ratio to the whole $(n-1)$-dimensional "area" of the hypersphere of that portion of it contained within a tube about C. If C is a great circle the solution of this problem is known; this is the case if $f(x, p)$ is a linear function of p. C is also a great circle if, when x is replaced by some function of a new variable ξ and p by some function of a variable π, $f(x, p)$ reduces to a linear function of π. In other cases C will be a curve other than a great circle. The determination whether an observed correlation is significant then requires the evaluation of the volume of a tube about C.

Similar considerations apply to other situations in which a parameter enters in a non-linear fashion. For example we may fit a regression equation of the form

$$Y = a + bf(x, p), \tag{1.2}$$

determining a, b and p so as to make $\Sigma(Y - y)^2$ a minimum. The reduction of its theory to that of (1.1) is effected in the following manner, which may easily be extended to other cases in which additional parameters enter into the regression equation linearly. Put

$$\bar{f}(p) = \Sigma f(x_a, p)/n, \qquad \bar{y} = \Sigma y_a/n$$
$$f'(x_a, p) = f(x_a, p) - \bar{f}(p), \qquad y'_a = y_a - \bar{y}.$$

The expression to be made a minimum,

$$\Sigma(Y_a - y_a)^2 = \Sigma\,\{a + bf\,(x_a, p) - y_a\}^2,$$

when expressed in terms of the quantities just introduced, reduces easily, since

$$\Sigma f'(x_a, p) = 0 = \Sigma y_a'$$

to

$$\Sigma\,\{bf'(x_a, p) - y_a'\}^2 + n\{a + b\bar{f}(p) - \bar{y}\}^2.$$

The minimizing of the first sum is of the same nature as in the case of the regression equation (1.1); when b and p are determined in this way, a is determined immediately so as to make the other term vanish. The distribution of the correlation between y and Y, eliminating the means in this case, is now determined by the volume of a tube on a unit hypersphere of $(n-2)$ dimensions in the flat space of $(n-1)$-dimensions whose equation is $\Sigma Y_a = 0$. The axis of the tube is the curve C whose equations are

$$\mu Y_a = f'(x_a, p) \qquad (a = 1, \cdots, n),$$

where μ is determined by the condition that $\Sigma Ya^2 = 1$.

The numerical process of fitting regression equations non-linear in parameters is considerably more laborious than in the linear case. It should be noticed that in all such problems, while transformations of parameters and also transformations of independent variates are permissible, it is not permissible to make a transformation of the dependent variate y without changing the hypotheses underlying the application of the method of least squares to the particular case. Thus, the common practice of taking logarithms of both sides of such a regression equation as $Y = be^{px}$ in order to reduce it to linear form leads to inexact results unless the errors in log y, rather than in y itself, can be regarded as normally distributed with a common mean and variance.

As generalizations of (1.1) and (1.2) we may consider regression equations involving two or more parameters in an essentially non-linear fashion. Outstanding among these are the harmonic of undetermined period,

(1.3) $$Y = a + b\cos(kx + \epsilon),$$

or more generally, a sum of such harmonics, and the logistic used to describe the growth of populations and of individual organisms,

(1.4) $$Y = \frac{b}{1 + me^{-ax}}.$$

It has not always been realized that periodogram analysis, at least in Schuster's original sense of fitting a harmonic of the form (1.3), is essentially a problem in least squares, and that the problem of significance is a special case of the general one of least squares. The only published exact test of significance is due to R. A. Fisher[3] and is predicated on the assumption that only those periods are to be considered that are submultiples of the whole range of observations available. Empirical scientists in search of periodicities in sunspots, light variation of stars, rainfall, and business fluctuations have however not confined themselves to such a limited set of trial periods. The procedure is rather to try a very large number of periods, perhaps greater than the number of observations, and select the one showing greatest intensity. This is virtually equivalent to solving by trial the normal equations corresponding to (1.3). The maximum intensity obtainable, divided by the mean square residual, will be a function of the correlation R between the observed values y and the values Y computed from the regression equation (1.3). The probability distribution of R in the absence of genuine periodicity, on the assumption of normally and independently distributed observations with a common mean and variance, may be found approximately for high values of R by the geometrical method. Indeed, by applying to (1.3) the same considerations by which the theory of (1.2) was reduced to that of (1.1), we arrive at the equations

$$(1.5) \qquad \mu Y_a = f'(x_a, k, \epsilon), \qquad (a = 1, \cdots, n)$$

satisfying the conditions $\Sigma Y_a = 0$ and $\Sigma Y_a^2 = 1$. The right-hand member is simply the difference between cos $(kx_a + \epsilon)$ and the mean of this expression for the various values of x_a (in applications, the times) corresponding to the observations. We may regard (1.5) as the equations of a two-dimensional surface with parameters k and ϵ, lying in the $(n-2)$-dimensional hypersphere whose equations are

$$\Sigma Y_a = 0, \quad \Sigma Y_a^2 = 1.$$

The probability of any particular value of R being exceeded is proportional to the volume of the hypersphere within a geodesic distance θ of this surface, where $R = \cos\theta$. If we confine the range of periods, that is, of values of k, so that the corresponding portion of the surface does not have too great curvatures, and if θ is not too great, it is evident that this probability will be exactly or approximately proportional to the area of the portion of the surface explored.

[3] "Tests of significance in harmonic analysis," *Proceedings of the Royal Society*, London, vol. 125 A (1929), pp. 54-59.

The method suggested by Fisher is equivalent to using a finite number, approximately $(n-3)/2$, of great circles on the hypersphere, at constant mutual geodesic distances from each other of a quarter of a great circle. Of these circles, the one nearest the sample point corresponds to the period of maximum intensity. The probability appropriate to a test of significance by this method is the ratio to the whole $(n-2)$-dimensional volume of the hypersphere of the sum of the volumes of all the tubes about the selected circles, of radii equal to the minimum distance from the sample point. The aggregate volume of all these tubes will evidently be less than the volume of the region within geodesic distance θ of the surface (1.5), which passes through the axes of the tubes. This merely means that the method allowing selection of any period in a continuous range gives a greater probability of a particular value of R being exceeded than does Fisher's method of confining attention to certain predetermined periods, as was to be expected. Also, if the critical probabilities are made equal for the two tests, some intensities significant by Fisher's method will not be significant by the method of continuous variability of period; while periods eliminated from consideration by Fisher's method will sometimes appear significant when they are admitted to consideration.

The logistic (1.4) may be dealt with similarly by finding the area of a surface of two dimensions in a hypersphere. But in this case the assumption of equal variances of the deviations for different values of y becomes questionable, and a transformation leading to a different form of the problem will usually be suggested by the application to be made. The logistic (1.4) satisfies the differential equation

$$\frac{d \log Y}{dx} = a - b'Y, \tag{1.6}$$

where $b' = a/b$. The assumptions ordinarily underlying the use of the logistic as a growth curve are more in keeping with the assumption of independence and uniform variance for the deviations between the two members of this differential equation than for the deviations between the members of (1.4). The parameters of (1.6) enter in a linear fashion, so that in its fitting classical methods are more appropriate than the relatively complex ones associated with the direct fitting of the integrated logistic equation (1.4), provided suitable estimates of the growth rate on the left of (1.6) are available. One method of dealing with this situation has been given by the author in an earlier paper.[4] An analogous method based on a difference equation

[4] "Differential equations subject to error, and population estimates," *Journal of the American Statistical Association*, vol. 22 (1927), pp. 283-314.

instead of a differential equation, had been given earlier by G. U. Yule;[5] it appears to be the better of the two from a practical standpoint when, as Yule assumes, the time intervals between observations are strictly uniform.

But in this paper we shall not deal further with problems involving more than a single non-linear parameter, nor shall we discuss the integrals whose evaluation is necessary for practical work with the examples indicated above. The subsequent sections are offered purely as contributions to geometry, except that the results of Section 3 are essential to the tests of significance just described. As is usual in differential geometry, we shall assume the functions involved to have in the neighborhoods concerned continuous finite derivatives of all orders essential to the argument. Latin indices will be used to indicate the values $1, 2, \cdots, n$, whereas Greek indices will take only the values $2, \cdots, n$ throughout the paper, except in § 3. Repetition of a Greek index within a term will denote summation from 2 to n; of a Latin index, summation from 1 to n.

2. Tubes in euclidean space. In terms of cartesian coördinates $x_1, x_2, \cdots, x_n$ let the curve C be defined by the equations

$$x_i = f_i(v_1), \tag{2.1}$$

where v_1 is the distance along the curve from some fixed point. We shall use primes to denote differentiation with respect to v_1. Denote by λ_{i1} the unit vector tangent to C, by λ_{i2} the unit first curvature vector of C, and by $\lambda_{i3} \cdots, \lambda_{in}$ a set of unit vectors orthogonal to each other and to λ_{i1} and λ_{i2}, so chosen that the determinant $|\lambda_{ij}| = +1$. Then $\lambda_{i1} = f_i'(v_1)$; also λ_{i1} equals its cofactor in the determinant.

Introducing curvilinear coördinates $v_1, v_2, \cdots, v_n$ by means of the relations

$$x_i = f_i(v_1) + v_\alpha \lambda_{i\alpha}(v_1), \tag{2.2}$$

where the last term represents a sum from 2 to n with respect to α, we have

$$J = \frac{\partial(x_1, \cdots, x_n)}{\partial(v_1, \cdots, v_n)} = \begin{vmatrix} \lambda_{11} + \lambda_{1\alpha}' v_\alpha & \lambda_{12} \cdots \lambda_{1n} \\ \cdots & \cdots \\ \lambda_{n1} + \lambda_{n\alpha}' v_\alpha & \lambda_{n2} \cdots \lambda_{nn} \end{vmatrix}.$$

Expanding with reference to the first column we obtain, since the cofactor of the j-th element in this column equals λ_{i1},

$$J = 1 + \lambda_{i1}\lambda_{i\alpha}' v_\alpha. \tag{2.3}$$

[5] *Journal of the Royal Statistical Society*, vol. 88 (1925), pp. 1-58.

Upon differentiating the orthogonality condition $\lambda_{i1}\lambda_{i\alpha} = 0$ we obtain:

$$\lambda_{i1}\lambda_{i\alpha}' + \lambda_{i1}'\lambda_{i\alpha} = 0. \tag{2.4}$$

The elementary relation between the principal normal, radius of first curvature ρ_1, and rate of change of the direction of the tangent may be written

$$\lambda_{i1}' = \frac{\lambda_{i2}}{\rho_1}. \tag{2.5}$$

Substituting this in (2.4), making use of the orthogonality of the vectors, and substituting the result in (2.3) gives:[6]

$$J = 1 - \frac{v_2}{\rho_1}. \tag{2.6}$$

It is clear that $v_2, \cdots, v_n$ are distances from the curve C in directions perpendicular to the tangent and to each other. A tubular hypersurface of radius θ therefore has the equation

$$v_2^2 + v_3^2 + \cdots + v_n^2 = \theta^2.$$

This may also be regarded as the equation of a hypersphere in space of $n-1$ dimensions. Upon integrating (2.6) with respect to $v_2, \cdots, v_n$ over the interior of this sphere, since the mean value of v_α is zero, we obtain merely the volume enclosed by the sphere, namely

$$\frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)}\theta^{n-1}. \tag{2.7}$$

Since this does not involve v_1, *the tubular volume corresponding to an arc is the product of* (2.7) *by the length of the arc.*

This result is exact, but takes no account of overlapping of the tube with itself. Overlapping may be of portions of the tube corresponding to non-consecutive arcs, or it may be a local phenomenon resulting from the curvature of the axial curve being excessive in relation to the radius of the tube. The first kind is not within the domain of differential geometry, and apparently nothing can be said about it without some further specializiation of the curve.

The second kind of overlapping, or kinking, will occur if and only if $J = 0$ at some point within the tube. Since $|v_2| < \theta$ within the tube, it is

---

[6] This method of evaluating J, which is simpler than my original reduction and does not require continuity of derivatives of the vectors of orders higher than the first, was pointed out by Dean L. P. Eisenhart, to whom I am indebted for reading this paper and making several suggestions.

evident from (2.6) that J vanishes within the tube if and only if $\theta > \rho_1$. Thus the condition for non-overlapping of the local sort is that the radius of the tube shall not exceed the radius of first curvature of the axial curve, regardless of curvatures of higher order.

3. Tubes on a hypersphere. In terms of cartesian coördinates $x_1, x_2 \cdots x_n$ in euclidean space of n dimensions we may write the equation of a unit hypersphere

$$\Sigma x^2 = 1. \tag{3.1}$$

In this section it will be convenient to use this type of notation, denoting summation from 1 to n by the sign Σ and frequently omitting subscripts. Let the curve C on this sphere be defined in terms of the arc length s by the n equations $x_i = x_i(s)$; let differentiation with respect to s be indicated by primes; and let $a_i = x_i'$. Then

$$\Sigma a^2 = 1, \tag{3.2}$$

and, by differentiation of (3.1),

$$\Sigma xa = 0. \tag{3.3}$$

Differentiating the last equation and using (3.2) we have:

$$\Sigma xa' = -1. \tag{3.4}$$

The radius ρ of first curvature relative to the euclidean space is given by

$$1/\rho^2 = \Sigma a'^2. \tag{3.5}$$

If the n new quantities $\xi_1, \cdots, \xi_n$ are subject to the three equations

$$\Sigma x\xi = 0, \qquad \Sigma a\xi = 0, \qquad \Sigma \xi^2 = 1, \tag{3.6}$$

it is evident that they have $n-3$ degrees of freedom for each value of s. Hence we may write them as functions,

$$\xi_i = \xi_i(s, \phi_1, \cdots \phi_{n-3}), \tag{3.7}$$

of forms to be specified later. Restricting the x_i to be cartesian coördinates of a point on the curve C, and therefore functions only of s, we shall use $y_1, \cdots, y_n$ as cartesian coördinates of a general point on the hypersphere, whose equation $\Sigma y^2 = 1$ is satisfied identically by the expressions

$$y_i = x_i \cos\theta + \xi_i \sin\theta, \tag{3.8}$$

because of (3.1) and (3.6). As curvilinear coördinates on the hypersphere we shall use s, θ, and the variables $\phi_1, \cdots \phi_{n-3}$ appearing in (3.7). Taking them in this order and using primes to denote *partial* differentiation with respect to s, we have as the matrix of coefficients of the linear element,

$$(3.9)\qquad \left\|\begin{matrix} \Sigma y'^2 & \Sigma y' \dfrac{\partial y}{\partial \theta} & \cdots & \\ \Sigma y' \dfrac{\partial y}{\partial \theta} & \Sigma \left(\dfrac{\partial y}{\partial \theta}\right)^2 & \cdots & \\ \sin\theta\, \Sigma y' \dfrac{\partial \xi}{\partial \phi_1} & \sin\theta\, \Sigma \dfrac{\partial y}{\partial \theta} \dfrac{\partial \xi}{\partial \phi_1} & \cdots & \sin^2\theta\, \Sigma \dfrac{\partial \xi}{\partial \phi_1} \dfrac{\partial \xi}{\partial \phi_{n-3}} \\ \cdot & \cdot & \cdot & \cdot \\ \sin\theta\, \Sigma y' \dfrac{\partial \xi}{\partial \phi_{n-3}} & \sin\theta\, \Sigma \dfrac{\partial y}{\partial \theta} \dfrac{\partial \xi}{\partial \phi_{n-3}} & \cdots & \sin^2\theta\, \Sigma \left(\dfrac{\partial \xi}{\partial \phi_{n-3}}\right)^2 \end{matrix}\right\|.$$

From (3.6) we have

$$(3.10)\qquad \Sigma x\xi' = 0,\quad \Sigma \xi\xi' = 0,\quad \Sigma x \frac{\partial \xi}{\partial \phi_\gamma} = 0,\quad \Sigma a \frac{\partial \xi}{\partial \phi_\gamma} = 0,\quad \Sigma \xi \frac{\partial \xi}{\partial \phi_\gamma} = 0 \qquad (\gamma = 1, \cdots, n-3).$$

From these and the preceding identities it is easy to see that the elements in the second row or second column of (3.9) are all zero, excepting the element in the intersection of the second row and column, which equals unity. This shows, by a well known theorem (*RG*,[7] p. 58) that θ measures the distance from C along geodesics of the hypersphere perpendicular to C.

Denoting by E the element in the upper left-hand corner of the matrix, we have from (3.8) and (3.2),

$$(3.11)\qquad E = \cos^2\theta + 2\cos\theta \sin\theta\, \Sigma a\xi' + \sin^2\theta\, \Sigma \xi'^2.$$

The other elements in the first column are given by

$$\sin\theta\, \Sigma(a\cos\theta + \xi' \sin\theta) \frac{\partial \xi}{\partial \phi_\gamma} = \sin^2\theta\, \Sigma \xi' \frac{\partial \xi}{\partial \phi_\gamma}$$

by (3.10). The element of $(n-1)$-dimensional volume on the hypersphere is the product of $ds\, d\theta d\phi_1, \cdots, d\phi_{n-3}$ by the square root of the determinant of (3.9). It therefore equals

$$(3.12)\qquad \sin^{n-3}\theta \sqrt{G}\, ds\, d\theta\, d\phi_1 \cdots d\phi_{n-3},$$

where

$$(3.13)\qquad G = \left|\begin{matrix} E & \sin\theta\, \Sigma \xi' \dfrac{\partial \xi}{\partial \phi_1} & \cdots & \sin\theta\, \Sigma \xi' \dfrac{\partial \xi}{\partial \phi_{n-3}} \\ \sin\theta\, \Sigma \xi' \dfrac{\partial \xi}{\partial \phi_1} & \Sigma \left(\dfrac{\partial \xi}{\partial \phi_1}\right)^2 & \cdots & \Sigma \dfrac{\partial \xi}{\partial \phi_1} \dfrac{\partial \xi}{\partial \phi_{n-3}} \\ \cdot & \cdot & \cdot & \cdot \\ \sin\theta\, \Sigma \xi' \dfrac{\partial \xi}{\partial \phi_{n-3}} & \Sigma \dfrac{\partial \xi}{\partial \phi_1} \dfrac{\partial \xi}{\partial \phi_{n-3}} & \cdots & \Sigma \left(\dfrac{\partial \xi}{\partial \phi_{n-3}}\right)^2 \end{matrix}\right|.$$

This determinant will be evaluated with the help of a special orthogonal

ennuple of vectors of the Schmidt type.[7] Denoting the i-th component of the j-th vector by λ_{ij} $(i, j, = 1, \cdots, n)$ we put $\lambda_{i1} = x_i$, and then define $\lambda_{i2}, \cdots, \lambda_{in}$ as linear functions of x_i and its successive derivatives with respect to s, such that each vector in the sequence involves a derivative of order higher by unity than the preceding vector, and such that the whole set is orthogonal and normal. Thus $\lambda_{i2} = \alpha_i$; and

$$\Sigma\lambda_{ik}\lambda_{im} = \delta_m{}^k \tag{3.14}$$

where $\delta_m{}^k$ is the Kronecker delta, equal to unity if $k = m$, and otherwise zero. The formulae *RG* (32.16) analogous to those of Frenet and Serret give in this case

$$\lambda_{ij}' = \frac{\lambda_{i,\,j+1}}{\rho_j} - \frac{\lambda_{i,\,j-1}}{\rho_{j-1}} \qquad (i, j = 1, \cdots, n), \tag{3.15}$$

where the convention is made that $1/\rho_0 = 1/\rho_n = 0$ and $\rho_1, \cdots, \rho_{n-1}$ are certain functions of s. (These λ's and ρ's are different from those of other sections of this paper.) Putting $j = 1$ in (3.15) and recalling that $\lambda_{i1}' = x_i' = a_i$, and that λ_{i2} also equals a_i, shows that $\rho_1 = 1$.

The quantities z_j defined by the orthogonal transformation

$$z_j = \Sigma\lambda_{ij}\xi_i \tag{3.16}$$

must, according to (3.6), satisfy the conditions

$$z_1 = z_2 = 0, \quad \Sigma z^2 = 1. \tag{3.17}$$

Multiplying (3.16) by λ_{mj} and summing with respect to j, using (3.14) and (3.17), and then changing m to i, gives

$$\xi_i = \Sigma\lambda_{ij}z_j = \lambda_{i3}z_3 + \cdots + \lambda_{in}z_n. \tag{3.18}$$

We may regard $z_3, \cdots, z_n$ as cartesian coördinates of a point independent of s on the sphere (3.17), whose intrinsic dimensionality is $n - 3$. We shall regard $\phi_1, \cdots, \phi_{n-3}$ as spherical polar coördinates on this sphere, thus specializing the arbitrary functions (3.7), which now take the form (3.18), where the λ's involve only s and the z's involve only the ϕ's. Putting

$$u_{j\gamma} = \frac{\partial z_j}{\partial \phi_\gamma} \tag{3.19}$$

we therefore obtain by differentiating (3.18),

$$\frac{\partial \xi}{\partial \phi_\gamma} = \Sigma\lambda_{ij}u_{j\gamma}, \tag{3.20}$$

---

[7] L. P. Eisenhart, Riemannian Geometry, Princeton, 1926, Sec. 32. We shall refer later to this treatise as RG.

and also, with the help of (3.15),

$$\xi_i' = \Sigma\lambda_{ij}'z_j = \Sigma\left(\frac{\lambda_{i,\,j+1}}{\rho_j} - \frac{\lambda_{i,\,j-1}}{\rho_{j-1}}\right)z_j. \tag{3.21}$$

Separating this into two summations, putting $j+1=k$ in the first and $j-1=k$ in the second, introducing the notation

$$\Delta_k = \frac{z_{k-1}}{\rho_{k-1}} - \frac{z_{k+1}}{\rho_k}, \tag{3.22}$$

and making use of the relations $z_1 = z_2 = 1/\rho_0 = 1/\rho_n = 0$, we find:

$$\xi_i' = \Sigma\lambda_{ik}\Delta_k. \tag{3.23}$$

From (3.22) it follows that

$$\Delta_1 = 0, \tag{3.24}$$

and from (3.23), (3.14), (3.24) and (3.20),

$$\Sigma\xi'^2 = \Sigma\Delta^2 = \Delta_2{}^2 + \Delta_3{}^2 + \cdots + \Delta_n{}^2, \tag{3.25}$$

$$\Sigma\alpha\xi' = \Sigma\lambda_{i2}\xi'_i = \Delta_2, \tag{3.26}$$

$$\Sigma\xi'\frac{\partial\xi}{\partial\phi_\gamma} = \Sigma u_{i\gamma}\Delta_i, \tag{3.27}$$

$$\Sigma\frac{\partial\xi}{\partial\phi_\gamma}\frac{\partial\xi}{\partial\phi_\delta} = \Sigma u_{i\gamma}u_{i\delta}. \tag{3.28}$$

From (3.11), (3.26) and (3.25),

$$E = (\cos\theta + \Delta_2\sin\theta)^2 + (\Delta_3{}^2 + \cdots + \Delta_n{}^2)\sin^2\theta. \tag{3.29}$$

Let the determinant in (3.13) be represented as the sum of two determinants, identical in all but the first column, in such a way that one of these determinants has as its first element $(\cos\theta + \Delta_2\sin\theta)^2$, and otherwise has zeros in the first column. We may thus write, with the help of (3.25) to (3.29) inclusive

$$G = F + H^2\sin^2\theta, \tag{3.30}$$

where

$$F = (\cos\theta + \Delta_2\sin\theta)^2\begin{vmatrix} \Sigma u_{i1}{}^2 \cdots & \\ \Sigma u_{i1}u_{i2} \cdots & \\ \cdot\quad\cdot\quad\cdot\quad\cdot & \cdot \\ & \cdots\Sigma u_{i}{}^2{}_{,\,n-3} \end{vmatrix}, \tag{3.31}$$

and

$$H = \begin{vmatrix} \Delta_3 \cdots \Delta_n \\ u_{31} \cdots u_{n1} \\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ u_{3,\,n-3} \cdots u_{n,\,n-3} \end{vmatrix}.$$

Differentiating the last of (3.6) gives $\Sigma\xi\xi' = 0$. If in this we substitute (3.18) and (3.23), and use (3.14), we obtain

$$\Sigma z_k \Delta_k = 0.$$

Differentiating $\Sigma z^2 = 1$ we have, from (3.19),

$$\Sigma z_k u_{k\gamma} = 0.$$

These equations establish a homogeneous linear relation among the columns of H. Hence $H = 0$, so that $G = F$ by (3.30). Denoting by G' the determinant in the right-hand member of (3.31), and noting from (3.22) that $\Delta_2 = - z_3/\rho_2$ we thus obtain the element of volume (3.12) in the form

$$(3.32) \qquad \sin^{n-3}\theta \, (\cos\theta - z_3 \sin\theta/\rho_2) \, \sqrt{G'} \, ds \, d\theta \, d\phi_1 \cdots d\phi_{n-3}.$$

The integral of $\sqrt{G'}\, d\phi_1 \cdots d\phi_{n-3}$ over the sphere is simply the $(n-3)$-dimensional volume of this unit sphere, namely

$$\frac{2\pi^{(n-2)/2}}{\Gamma\left(\frac{n-2}{2}\right)}.$$

In integrating (3.32), the integral resulting from the second term in the parenthesis vanishes because z_3 is measured perpendicularly from a diametral plane of the sphere, and so has a mean value zero. Integrating also with respect to θ and s, we have the simple result:

The volume enclosed by a tube of geodesic radius θ on a hypersphere having intrinsically $n-1$ dimensions is the product of the length of the axial curve by

$$\frac{\pi^{(n-2)/2}\sin^{n-2}\theta}{\Gamma(n/2)}.$$

Local self-overlapping will exist if (3.32) vanishes within the tube. This will occur if and only if $\tan\theta > \rho_2$, where θ is the geodesic radius. To evaluate ρ_2 we first put $j = 2$ in (3.15) and deduce; since $\lambda_{i1} = x_i$, $\lambda_{i2} = \alpha_i$ and $\rho_1 = 1$, that

$$\lambda_{i3} = \rho_2\,(x_i + \alpha'_i).$$

Squaring, summing with respect to i, and using (3.14), (3.1), (3.4) and (3.5), we find

$$\rho_2 = \frac{\rho}{\sqrt{1-\rho^2}}.$$

The condition $\tan\theta \leq \rho_2$ for absence of local self-overlapping is therefore equivalent to

$$\sin\theta \leq \rho.$$

This condition is also expressed by the statement that the geodesic radius of the tube must not exceed the maximum radius of geodesic curvature of C if there is to be no local overlapping.

As an application, we observe that in fitting the regression equation

$$Y = be^{px}$$

we obtain a curve which, for $p = \pm \infty$, has ends. At these ends, the radius of curvature becomes zero. Consequently, if the foregoing proposition regarding the volume of a tube is to be applied to evaluate the goodness of fit, it is necessary either to confine attention to values of $|p|$ less than some upper limit, or to make a special study of the volume in neighborhoods of the ends of the curve. In the former case the volumes of hemispherical caps over the ends should be added to that enclosed by the tube in determining the relevant probability.

4. An orthogonality property.[8] The following theorem concerns geodesic spheres in an arbitrary Riemannian space; it reduces to one of Gauss when the space is of two dimensions, and may be proved in a somewhat similar manner:[9]

The geodesic sphere defined as the locus of points at a fixed geodesic distance from a point O is perpendicular to the geodesics through O.

The differential equations of the geodesics in terms of the arc length s are (RG, (17.8)):

$$\frac{d^2x^i}{ds^2} + \left\{ {i \atop j\,k} \right\} \frac{dx^j}{ds}\frac{dx^k}{ds} = 0. \tag{4.1}$$

Taking the geodesics through O as the coördinate lines along which $x^1 = s$ is the distance from O and the other coördinates $x^2, \cdots, x^n$ are constant, we have

$$\frac{dx^j}{ds} = \delta_1{}^j.$$

Substituting this in (4.1) gives

$$\left\{ {i \atop 1\,1} \right\} = 0. \qquad (i = 1, 2, \cdots, n).$$

Hence $[1\,1, \alpha] = 0$. Since the choice of coördinates implies that $g_{11} = 1$ identically, it then follows that

$$\frac{\partial g_{1\alpha}}{\partial x^1} = 0. \tag{4.2}$$

Consider also another system of coördinates $y^1, \cdots, y^n$ such that $y^i = \xi^i x^1$, where the ξ^i are any functions of $x^2, \cdots, x^n$ have finite derivatives in a neigh-

[8] In this and the following sections the notation is throughout that of RG. Latin indices will in all cases vary from 1 to n, Greek indices from 2 to n.

[9] L. P. Eisenhart, Differential Geometry, Boston, 1909, p. 207.

borhood of O. Denoting by g'_{jk} the components of the distance tensor in this coordinate system, we have

$$g_{1a} = g'_{jk} \frac{\partial y^j}{\partial x^1} \frac{\partial y^k}{\partial x^a} = g'_{jk} \xi^j \frac{\partial \xi^k}{\partial x^a} x^1.$$

Thus at O, where $x^1 = 0$, $g_{ia} = 0$. Since (4.2) shows that g_{ia} is independent of x^1, it follows that $g_{ia} = 0$ everywhere. This proves the theorem. Another proof of this theorem, based on the transversality condition of the calculus of variations could also be given.

5. Spheres in a general curved space. Let $x^1, \cdots, x^n$ be normal coördinates with origin at the center 0 of a geodesic sphere of radius θ. The element of volume is

$$\sqrt{g}\, dx^1 \cdots dx^n, \tag{5.1}$$

where g is the determinant of the distance tensor g_{ij}. For a point on the sphere let $\xi^1, \cdots, \xi^n$ be defined by the equations

$$x^i = \xi^i \theta, \tag{5.2}$$

which are also the equations of the geodesics through 0 if θ is regarded as a parameter and the ξ^i as constants. The ξ^i may be regarded as cartesian coördinates of a point on a unit sphere in euclidean space of n dimensions. Let $M_n\{\phi\}$ denote the mean value over this $(n-1)$-dimensional sphere of a function ϕ. Obviously $M_n\{1\} = 1$, and $M_n\{\xi^i\} = 0$. From considerations of symmetry it is further obvious that the mean value of the product of any powers of the ξ^i vanishes unless each of the ξ^i enters into the product with an even exponent. By integration with respect to spherical polar coördinates, or in various other ways, it is easy to establish that

$$M_n\{(\xi^i)^2\} = \frac{1}{n}, \; M_n\{(\xi^i)^4\} = \frac{3}{n(n+2)}, \tag{5.3}$$

and at the same time, that the $(n-1)$-dimensional volume itself is

$$A_{n-1} = \frac{2\pi^{(n/2)}}{\Gamma\left(\frac{n}{2}\right)}. \tag{5.4}$$

At the origin of normal coördinates, $g_{ij} = \delta^i_j$ and consequently $g = 1$. Upon expanding $\sqrt{g}$ in a series of powers of the x^i and substituting from (5.2) we have therefore

$$\sqrt{g} = 1 + \left[\frac{\partial \sqrt{g}}{\partial x^i}\right]_0 \xi^i \theta + \left[\frac{\partial^2 \sqrt{g}}{\partial x^i\, \partial x^j}\right]_0 \xi^i \xi^j \frac{\theta^2}{2} + \cdots. \tag{5.5}$$

The $(n-1)$-dimensional volume of the sphere is found by integrating this

expression over the unit euclidean sphere and multiplying by θ^{n-1}. In this process all terms of odd order in θ vanish, since they are multiplied by odd numbers of the ξ^i. We thus obtain

$$(5.6)\quad A_{n-1}\,\theta^{n-1}\Big\{1+\frac{\theta^2}{2n}\Sigma\Big[\frac{\partial^2\sqrt{g}}{(\partial x^i)^2}\Big]_0 + \frac{\theta^4}{24n}\frac{3}{(n+2)}\Big(\Sigma\Big[\frac{\partial^4\sqrt{g}}{(\partial x^i)^4}\Big]_0 + 2\sum_{i<j}\Big[\frac{\partial^4\sqrt{g}}{(\partial x^i)^2(\partial x^j)^2}\Big]_0\Big)+\cdots\Big\}.$$

In terms of normal coördinates we have from RG (18.8) and (18.9)

$$(5.7)\qquad \left\{ {m \atop i\ j} \right\}_0 = 0,$$

$$(5.8)\qquad \left(\frac{\partial\left\{ {m \atop i\ j} \right\}}{\partial x^k} + \frac{\partial\left\{ {m \atop k\ i} \right\}}{\partial x^j} + \frac{\partial\left\{ {m \atop j\ k} \right\}}{\partial x^i}\right)_0 = 0,$$

and further identities which we shall not use here, since we shall limit our consideration of the series (5.6) to evaluating the second term in invariant form. The Ricci tensor, RG (8.14), is

$$(5.9)\quad R_{ij} = \frac{\partial^2\log\sqrt{g}}{\partial x^i\,\partial x^j} - \frac{\partial\left\{ {k \atop i\ j} \right\}}{\partial x^k} + \left\{ {m \atop i\ k} \right\}\left\{ {k \atop m\ j} \right\} - \left\{ {m \atop i\ j} \right\}\frac{\partial\log\sqrt{g}}{\partial x^m}.$$

From RG (7.9), namely,

$$(5.10)\qquad \frac{\partial\log\sqrt{g}}{\partial x^i} = \left\{ {k \atop k\ i} \right\},$$

we have

$$(5.11)\qquad \frac{\partial^2\log\sqrt{g}}{\partial x^i\,\partial x^j} = \frac{\partial\left\{ {k \atop k\ i} \right\}}{\partial x^j} = \frac{\partial\left\{ {k \atop k\ j} \right\}}{\partial x^i}.$$

Putting $m=k$ in (5.8), summing for k, and using (5.11), gives

$$(5.12)\qquad \left(\frac{\partial\left\{ {k \atop i\ j} \right\}}{\partial x^k}\right)_0 = -2\Big[\frac{\partial^2\log\sqrt{g}}{\partial x^i\,\partial x^j}\Big]_0.$$

From (5.9), (5.12) and (5.7),

$$(5.13)\qquad [R_{ij}]_0 = 3\Big[\frac{\partial^2\log\sqrt{g}}{\partial x^i\,\partial x^j}\Big]_0 = 3\Big[\frac{\partial^2\sqrt{g}}{\partial x^i\,\partial x^j}\Big]_0,$$

since $(\sqrt{g})_0=1$ and, by (5.10) and (5.7), $(\partial\sqrt{g}/\partial x^i)_0=0$. The scalar curvature of the space is defined as

$$(5.14)\qquad R = g^{ij}R_{ij}.$$

Since $(g^{ij})_0 = \delta_j^i$, the value taken by R at O is $\Sigma(R_{ii})_0$. Combining this result with (5.13) and (5.6), we have for the $(n-1)$-dimensional volume,

$$A_{n-1}\theta^{n-1}\left\{1+\frac{R_0\theta^2}{6n}+\cdots\right\}, \tag{5.15}$$

where R_0 is the scalar curvature at the center. Integrating with respect to θ we have as the n-dimensional volume enclosed by the sphere,

$$\frac{\pi^{(n/2)}}{\Gamma\left(\frac{n+2}{2}\right)}\theta^n\left\{1+\frac{R_0\theta^2}{6(n+2)}+\cdots\right\}. \tag{5.16}$$

The case $n=2$ of these results is due to Bertrand and Diguet. Their results are obtained by putting $n=2$, $R_0=-2K$, where K is the Gaussian curvature of the surface, in (5.15) and (5.16).

6. Tubes in a general curved space. Referring once more to RG § 32, we make use of the special Schmidt orthogonal ennuple for which $\lambda^i_{1/} = dx^i/ds$ is the direction of the tangent to the curve C. Hence from the generalized Frenet formulae RG (32.16) for spaces of positive definite distance element,

$$\lambda^i_{p/,j}\,\lambda^j_{1/} = \frac{\lambda^i_{p+1/}}{\rho_p} - \frac{\lambda^i_{p-1/}}{\rho_{p-1}}, \tag{6.1}$$

where $\rho_1, \cdots, \rho_{n-1}$ are the successive curvatures of C, and $1/\rho_0 = 1/\rho_n = 0$. These equations hold in every coördinate system.

We shall denote by x^1 the arc distance along C from some fixed point, and define x^1 at other points of the space by the condition that it shall be constant on every geodesic perpendicular to C. We restrict attention to a region such that no two geodesics normal to C at points of the region meet again within the region, and every point of the region lies on such a geodesic. A point P of this region then lies on a unique geodesic normal to the curve C; let Q be the point at which this geodesic meets C, and let s be the geodesic distance QP. Let ξ^α be the cosine of the angle at Q between the direction QP and the vector $\lambda^i_{\alpha/}$ of the orthogonal set described above. We define[10] the α-th coördinate of P as $x^\alpha = \xi^\alpha s$. The equations of the geodesics QP in terms of the arc s as parameter are therefore

$$x^1 = \text{constant}, \qquad x^\alpha = \xi^\alpha s. \tag{6.2}$$

Since the vectors of the Schmidt ennuple are mutually orthogonal and are tangent at points of C to the coördinate lines (i. e., the curves along each of

---

[10] **It is important to bear in mind that in this section Greek indices take only the values 2, . . ., n, while Latin indices vary from 1 to n.**

which only one coördinate varies), we have at points of C:

$$g_{ij} = \delta_j^{\,i} = g^{ij}, \qquad g = 1. \tag{6.3}$$

If we substitute the equations (6. 2) in the differential equations (4. 1) of geodesics, we obtain:

$$\left\{ \begin{matrix} i \\ \alpha\beta \end{matrix} \right\} \xi^\alpha \xi^\beta = 0.$$

These equations are valid throughout the region, though both the Christoffel symbols and the ξ^α depend on the point P of evaluation. But at a point of C the Christoffel symbols take on definite values, because of the continuity assumed at the end of § 1, while the last equations hold when any numbers whatever are substituted for the ξ^α. Since a quadratic form can vanish for all sets of values of the variables only if all the coefficients vanish, we must have at every point of C,

$$\left\{ \begin{matrix} i \\ \alpha\beta \end{matrix} \right\} = 0. \tag{6.4}$$

If we differentiate (4. 1) with respect to s and substitute (6. 2) we obtain similarly the equations, valid at all points of C,

$$\frac{\partial \left\{ \begin{matrix} i \\ \alpha\beta \end{matrix} \right\}}{\partial x^\gamma} + \frac{\partial \left\{ \begin{matrix} i \\ \gamma\alpha \end{matrix} \right\}}{\partial x^\beta} + \frac{\partial \left\{ \begin{matrix} i \\ \beta\gamma \end{matrix} \right\}}{\partial x^\alpha} = 0. \tag{6.5}$$

In what follows it will be understood that the expressions considered are evaluated on C. Since (6. 4) holds at all points of C, irrespectively of the value of x^1, the derivative of the left member with respect to x^1 vanishes on the curve. In particular,

$$\frac{\partial \left\{ \begin{matrix} 1 \\ \alpha\beta \end{matrix} \right\}}{\partial x^1} = 0. \tag{6.6}$$

From the definition of the coördinates above it follows that the direction of the coördinate line along which only x^p varies must at points of C coincide with that of the unit vector $\lambda^i{}_{p/}$. Hence at such a point this coördinate line must satisfy

$$\frac{dx^i}{ds} = \lambda^i{}_{p/}.$$

But such a line must also by its very nature satisfy

$$\frac{dx^i}{ds} = \delta_p^{\,i}.$$

Therefore $\lambda^i_{p/} = \delta^i_p$ on C. Elsewhere $\lambda^i_{p/}$ has not been defined, but for convenience we define

$$\lambda^i_{p/} = \delta^i_p \tag{6.7}$$

at all other points of the space. The covariant derivative of this contravariant vector is by definition

$$\lambda^i_{p/,j} = \frac{\partial \lambda^i_{p/}}{\partial x^j} + \lambda_{p/}{}^k \left\{ \begin{matrix} i \\ k\, j \end{matrix} \right\}.$$

With (6.7) this gives

$$\lambda^i_{p/,j} = \left\{ \begin{matrix} i \\ p\, j \end{matrix} \right\}. \tag{6.8}$$

Substituting (6.7) and (6.8) in (6.1) we have, at points of C,

$$\left\{ \begin{matrix} i \\ p\, 1 \end{matrix} \right\} = \frac{\delta^i_{p+1}}{\rho_p} - \frac{\delta^i_{p-1}}{\rho_{p-1}}. \tag{6.9}$$

In particular,

$$\left\{ \begin{matrix} 1 \\ a\, 1 \end{matrix} \right\} = - \frac{\delta^2_a}{\rho_1}. \tag{6.10}$$

The components of the Ricci tensor (5.9) with subscripts $2, \cdots, n$ simplify on account of (6.4) and (6.6) to the form

$$R_{a\beta} = \frac{\partial^2 \log \sqrt{g}}{\partial x^a\, \partial x^\beta} - \frac{\partial \left\{ \begin{matrix} \gamma \\ a\, \beta \end{matrix} \right\}}{\partial x^\gamma} + \left\{ \begin{matrix} 1 \\ a\, 1 \end{matrix} \right\} \left\{ \begin{matrix} 1 \\ 1\, \beta \end{matrix} \right\}. \tag{6.11}$$

The scalar curvature (5.14) may with the help of (6.3), (6.11) and (6.10) be expressed in the form

$$\begin{aligned} R &= R_{11} + g^{a\beta} R_{a\beta} \\ &= R_{11} + g^{a\beta} \left(\frac{\partial^2 \log \sqrt{g}}{\partial x^a\, \partial x^\beta} - \frac{\partial \left\{ \begin{matrix} \gamma \\ a\, \beta \end{matrix} \right\}}{\partial x^\gamma} \right) + \frac{1}{\rho_1{}^2}. \end{aligned} \tag{6.12}$$

The *mean curvature* of the space at a point with respect to the direction $\lambda^i_{1/}$ of the curve is defined as

$$R' = R_{ij} \lambda^i_{1/} \lambda^j_{1/}.$$

In view of (6.7) this gives

$$R' = R_{11}. \tag{6.13}$$

The mean curvature with respect to any direction has the following geometrical meaning (*RG*, p. 113). With each of $n-1$ directions orthogonal to the given direction and to each other, the given direction determines a pencil of

geodesics forming a surface. The sum of the Gaussian curvatures of these $n-1$ surfaces is the negative of the mean curvature of the space for the given direction. From this we derive a geometrical interpretation of the scalar curvature. Since in normal coördinates

$$R = R_{11} + R_{22} + \cdots + R_{nn},$$

and since each term on the right is, like (6.13), the mean curvature with respect to a particular one of an orthogonal ennuple of directions, it follows that $-R$ is twice the sum of the Gaussian curvatures of all the $n(n-1)/2$ geodesic surfaces determined by these directions. If further we denote by S the scalar curvature of the hypersurface $x^1 =$ constant, we have from this interpretation that $-S$ is twice the sum of the Gaussian curvatures of those geodesic surfaces determined by the ennuple which lie in the hypersurface. From this it follows that

$$R = S + 2R'. \tag{6.14}$$

With reference to the hypersurface $x^1 =$ constant the components $g_{\alpha\beta}$ of the distance tensor have the same values as for the n-space. The same is therefore true of those Christoffel symbols whose indices have the values $2, \cdots, n$, and of the derivatives with respect to $x^2, \cdots, x^n$ of these symbols. We shall denote by h the $(n-1)$-rowed determinant of the $g_{\alpha\beta}$, and by $S_{\alpha\beta}$ the Ricci tensor of the hypersurface where it is pierced by C. Similarly to (5.11) we have:

$$\frac{\partial^2 \log \sqrt{h}}{\partial x^\alpha\, \partial x^\beta} = \frac{\partial \left\{ \begin{matrix} \gamma \\ \gamma\, \alpha \end{matrix} \right\}}{\partial x^\beta} = \frac{\partial \left\{ \begin{matrix} \gamma \\ \gamma\, \beta \end{matrix} \right\}}{\partial x^\alpha}. \tag{6.15}$$

If in (6.5) we replace i by γ and then sum with respect to γ from 2 to n we obtain with the help of (6.15)

$$\frac{\partial \left\{ \begin{matrix} \gamma \\ \alpha\, \beta \end{matrix} \right\}}{\partial x^\gamma} = -2\, \frac{\partial^2 \log \sqrt{h}}{\partial x^\alpha\, \partial x^\beta}. \tag{6.16}$$

The Ricci tensor $S_{\alpha\beta}$ is obtained from the right-hand member of (5.9) by replacing i, j, k, m respectively by $\alpha, \beta, \gamma, \delta$ and using h in places of g. With (6.4) and (6.16) this gives

$$S_{\alpha\beta} = -\frac{3}{2}\, \frac{\partial \left\{ \begin{matrix} \gamma \\ \alpha\, \beta \end{matrix} \right\}}{\partial x^\gamma}. \tag{6.17}$$

Since $S = g^{\alpha\beta} S_{\alpha\beta}$ we have from (6.17) and (6.14),

$$g^{\alpha\beta}\frac{\partial\left\{\begin{matrix}\gamma\\ \alpha\beta\end{matrix}\right\}}{\partial x^\gamma}=-\frac{2}{3}R+\frac{4}{3}R'. \tag{6.18}$$

Substituting this and (6.13) in (6.12) gives, after rearrangement,

$$g^{\alpha\beta}\frac{\partial^2\log\sqrt{g}}{\partial x^\alpha\,\partial x^\beta}=\frac{R+R'}{3}-\frac{1}{\rho_1{}^2}. \tag{6.19}$$

From (5.10), (6.10), (6.4) and (6.3) it follows that

$$\frac{\partial\sqrt{g}}{\partial x^\alpha}=-\frac{\delta_\alpha{}^2}{\rho_1}.$$

Hence (6.19) gives

$$g^{\alpha\beta}\frac{\partial^2\sqrt{g}}{\partial x^\alpha\,\partial x^\beta}=\frac{R+R'}{3}. \tag{6.20}$$

For a fixed value of x^1 we may expand $\sqrt{g}$ in a series of powers of $x^2,\cdots,x^n$, replace x^α by $\xi^\alpha\theta$ to obtain a series resembling (5.5) but with Latin indices replaced by Greek, and then integrate over the $(n-2)$-dimensional volume of the sphere $\Sigma(\xi^\alpha)^2=1$. This gives for the volume element of a tube $\sqrt{G}dx^1\,d\theta$, where

$$\begin{aligned}\sqrt{G}&=A_{n-2}\,\theta^{n-2}M_{n-1}\{\sqrt{g}\}\\ &=A_{n-2}\,\theta^{n-2}\left[(\sqrt{g})_C+\theta\left(\frac{\partial\sqrt{g}}{\partial x^\alpha}\right)_C M_{n-1}\{\xi^\alpha\}+\frac{\theta^2}{2}\left(\frac{\partial^2\sqrt{g}}{\partial x^\alpha\,\partial x^\beta}\right)_C M_{n-1}\{\xi^\alpha\xi^\beta\}+\cdots\right].\end{aligned}$$

The symmetry considerations of § 5 together with (5.3), (5.4), (6.3) and (6.20) reduce this to

$$\sqrt{G}=\frac{2\pi^{(n-1)/2}\theta^{n-2}}{\Gamma\left(\dfrac{n-1}{2}\right)}\left[1+(R+R')\frac{\theta^2}{6(n-1)}+\cdots\right].$$

Integrating with respect to θ gives as the volume element of a tube of radius θ,

$$\frac{\pi^{(n-1)/2}}{\Gamma\left(\dfrac{n+1}{2}\right)}\theta^{n-1}\left[1+\frac{(R+R')\,\theta^2}{6\,(n+1)}+\cdots\right]dx^1.$$

It might have been thought on the basis of geometric visualization that this result could have been obtained from (5.16) by replacing n by $n-1$ and R_0 by S. But R_0 must be replaced, not by S, but by $R+R'=S+3R'$.

By an extension of the foregoing procedures it seems likely that a fairly straightforward calculation would give the terms of these series, and of the

corresponding series (5.15) and (5.16), to any required degree. What is required is to express the symmetrical sums of higher derivativs of $\sqrt{g}$ in terms of invariants by formulae analogous to (6.20). Invariants available for the purpose are the higher covariant derivatives of the right-hand member of (6.20), and the contracted covariant derivatives of the Ricci tensor. It is a question of some interest whether the volume element of the tube also involves the various radii of curvature ρ_k of C. It does not involve them in either of the two cases we have examined fully, those of euclidean and of spherical space. If they do enter in other cases, these will doubtless call for the use of (6.9), a formula whose use could have been avoided in obtaining only the terms found above.

The conditions for non-overlapping in terms of the radius of first curvature found for euclidean and spherical spaces do not seem capable of generalization to arbitrary spaces. The condition applicable instead is that the radius of the tube shall be so small that no two geodesics through the curve and perpendicular to it shall meet again within the tube.

COLUMBIA UNIVERSITY.

ON THE VOLUME OF TUBES.*

By HERMANN WEYL.

1. The problem. In a lecture before the Mathematics Club at Princeton last year Professor Hotelling stated the following geometric problem [1] as one of primary importance for certain statistical investigations:

Let there be given in the n-dimensional Euclidean space E_n or spherical space S_n a closed ν-dimensional manifold C_ν. The solid spheres of given radius a around all the points of C_ν cover a certain part $C_\nu(a)$ of the embedding space E_n or S_n, the volume $V(a)$ of which is to be determined. We call $C_\nu(a)$ an (n,ν)-tube (of radius a around C_ν).

For small values of a one will have in the first approximation

$$V(a) = \Omega_m a^m \cdot k_0,$$

where $\Omega_m a^m$ is the volume of the solid m-dimensional sphere

$$\sigma_m(a): \quad t_1^2 + \cdots + t_m^2 \leqq a^2 \tag{1}$$

$(m = n - \nu)$, and k_0 the area of the "surface" C_ν. Professor Hotelling showed that this formula is exact in E_n and a similar formula prevails in S_n, *for* $\nu = 1$. I shall here treat the problem for higher dimensionalities ν. The result in E_n is a formula consisting of $1 + [\frac{1}{2}\nu]$ terms, of the following type (§ 3):

$$V(a) = \Omega_m \cdot \sum_e \frac{a^{m+e}}{(m+2)(m+4)\cdots(m+e)} k_e \qquad (e \text{ even}, \ 0 \leqq e \leqq \nu), \tag{2}$$

where k_e is a certain integral invariant of the surface C_ν determined by the intrinsic metric nature of C_ν only, and thus independent of its embedding in E_n. I shall express these invariants (§ 4) in terms of the Riemannian tensor of C_ν. An analogous result is obtained for S_n.

2. The fundamental formulas for the volume of tubes. If an n-dimensional manifold M_n consisting of points u and locally referred to

* Received October 14, 1938.

[1] See his paper "Tubes and spheres in n-spaces, and a class of statistical problems" which precedes this article in this Journal, pp. 440-460.

parameters $u^1, \cdots, u^n$ is mapped upon the Euclidean space E_n with the coördinates $(x_1, \cdots, x_n) = \mathfrak{r}$,

$$(3) \qquad \mathfrak{r} = \mathfrak{r}(u) = \mathfrak{r}(u^1, \cdots, u^n),$$

then the volume V of the image of M_n in E_n may be computed by means of the formula

$$(4) \qquad V = \int [\mathfrak{r}_1 \cdots \mathfrak{r}_n] du^1 \cdots du^n,$$

where $[\mathfrak{r}_1 \cdots \mathfrak{r}_n]$ designates the determinant of the n columns $\mathfrak{r}_i$, each consisting of the components of the vector

$$\mathfrak{r}_i = \partial \mathfrak{r} / \partial u^i.$$

This formula takes account of the $\pm$ orientation and multiplicity with which the mapping $u \rightarrow \mathfrak{r}$ covers the several parts of E_n. The covering will be locally a one-to-one mapping without folds and ramifications wherever $[\mathfrak{r}_1 \cdots \mathfrak{r}_n] > 0$. But even if this condition is satisfied everywhere, multiple covering might occur. This question is essentially one of topological rather than differential geometric nature. It is with this reservation in mind that in the following we apply formula (4).

When dealing with the spherical space S_n we employ homogeneous coördinates $(x_0, x_1, \cdots, x_n) = \mathfrak{r}$, the set ρx_i meaning the same point as x_i, whatever the factor $\rho \neq 0$. Sometimes we use the normalization

$$\mathfrak{r}^2 = x_0^2 + x_1^2 + \cdots + x_n^2 = 1.$$

S_n then appears as the unit sphere in the Euclidean E_{n+1}. (4) must be replaced by the formula

$$(5) \qquad V = \int \frac{[\mathfrak{r}\mathfrak{r}_1 \cdots \mathfrak{r}_n]}{(\mathfrak{r}^2)^{(n+1)/2}} du^1 \cdots du^n$$

as one easily verifies by observing the following facts: (1) the integrand is orthogonally invariant; (2) it is not affected by the gauge factor $\rho = \rho(u)$ because

$$(\rho \mathfrak{r})_i = \rho \cdot \mathfrak{r}_i + \frac{\partial \rho}{\partial u^i} \cdot \mathfrak{r};$$

(3) at the point $\mathfrak{r} = (1, 0, \cdots, 0)$ the integrand reduces to the "Euclidean" value

$$\begin{vmatrix} \dfrac{\partial x_1}{\partial u^1}, & \cdots, & \dfrac{\partial x_n}{\partial u^1} \\ \cdot & \cdot & \cdot \\ \dfrac{\partial x_1}{\partial u^n}, & \cdots, & \dfrac{\partial x_n}{\partial u^n} \end{vmatrix}.$$

After these preliminary remarks I now turn to our problem in E_n. Let a piece of the ν-dimensional manifold C_ν be given in the Gaussian representation

$$\mathfrak{r} = \mathfrak{r}(u^1 \cdots u^\nu). \tag{6}$$

At each point we can determine $m = n - \nu$ normal vectors $\mathfrak{n} = \mathfrak{n}(1), \cdots, \mathfrak{n}(m)$ satisfying the equations

$$\mathfrak{r}_\alpha \cdot \mathfrak{n} = 0 \qquad (\alpha = 1, \cdots, \nu)$$

which are mutually normalized by

$$\mathfrak{n}(p) \cdot \mathfrak{n}(q) = \delta_{pq} \qquad (p, q = 1, \cdots, m).$$

$\mathfrak{r}_\alpha$ is the derivative $\partial\mathfrak{r}/\partial u^\alpha$. In using the radius vector (6) and these normals, the part $C_\nu(a)$ of the space covered by the spheres of radius a around the points of C_ν allows the representation

$$\mathfrak{x} = \mathfrak{r} + t_1\mathfrak{n}(1) + \cdots + t_m\mathfrak{n}(m), \qquad (t_1{}^2 + \cdots + t_m{}^2 \leqq a^2), \tag{7}$$

in terms of the parameters $u^1, \cdots, u^\nu, t_1, \cdots, t_m$. Hence its volume $V(a)$ is the integral

$$\int [\mathfrak{x}_1, \cdots, \mathfrak{x}_\nu, \mathfrak{n}(1), \cdots, \mathfrak{n}(m)]\, dt_1 \cdots dt_m du^1 \cdots du^\nu. \tag{8}$$

Following Gauss we describe the surface C_ν embedded in E_n by its metric ground form

$$(d\mathfrak{r})^2 = \sum_{\alpha,\beta} g_{\alpha\beta} du^\alpha du^\beta, \qquad (g_{\alpha\beta} = \mathfrak{r}_\alpha \cdot \mathfrak{r}_\beta)$$

together with the linear pencil of the second fundamental forms

$$-\sum_{p=1}^{m} \{t_p \sum_{\alpha,\beta} G_{\alpha\beta}(p)\, du^\alpha du^\beta\},$$

which is the scalar product of

$$d^2\mathfrak{r} = \sum_{\alpha,\beta} \mathfrak{r}_{\alpha\beta} du^\alpha du^\beta$$

with an arbitrary normal $\mathfrak{n} = t_1\mathfrak{n}(1) + \cdots + t_m\mathfrak{n}(m)$.

$$G_{\alpha\beta}(p) = G_{\beta\alpha}(p) = -\mathfrak{r}_{\alpha\beta} \cdot \mathfrak{n}(p) = \mathfrak{r}_\alpha \cdot \mathfrak{n}_\beta(p).$$

A Greek subscript α attached to the vectors $\mathfrak{r}$, $\mathfrak{n}$ and $\mathfrak{x}$ always denotes differentiation by u^α.

Each vector at the point u of C_ν is a linear combination of the basic vectors $\mathfrak{r}_a$, $\mathfrak{n}(p)$. On applying this remark to $\mathfrak{n}_a(p)$ we set

$$\mathfrak{n}_a(p) = \sum_\beta G_a{}^\beta(p) \cdot \mathfrak{r}_\beta + \cdots,$$

where . . . indicates a linear combination of the normal vectors $\mathfrak{n}(p)$. By scalar multiplication with $\mathfrak{r}_\beta$ one finds

$$G_{a\beta}(p) = \sum_\lambda g_{a\lambda} G_\beta{}^\lambda(p).$$

From (7) one infers that

$$\mathfrak{x}_a = \sum_\beta \{\delta_a{}^\beta + \sum_{p=1}^m t_p G_a{}^\beta(p)\}\mathfrak{r}_\beta + \cdots.$$

Therefore the integrand in (8)

$$= \det \{\delta_a{}^\beta + \sum_p t_p G_a{}^\beta(p)\} \cdot [\mathfrak{r}_1 \cdots \mathfrak{r}_\nu \mathfrak{n}(1) \cdots \mathfrak{n}(m)].$$

Because of the general identity

$$[\mathfrak{a}_1 \cdots \mathfrak{a}_n]^2 = \det(\mathfrak{a}_i \mathfrak{a}_k),$$

$$[\mathfrak{r}_1 \cdots \mathfrak{r}_\nu \mathfrak{n}(1) \cdots \mathfrak{n}(m)]^2 = |\, g_{a\beta} \,|,$$

and considering that

$$ds = |\, g_{a\beta} \,|^{\frac{1}{2}} du^1 \cdots du^\nu$$

is the area element of C_ν, one arrives at the fundamental formula

$$V(a) = \int_{C_\nu} \left\{ \int \cdots \int_{(t_1^2 + \ldots + t_m^2 \leqq a^2)} |\, \delta_a{}^\beta + \sum_p t_p G_a{}^\beta(p) \,|\, dt_1 \cdots dt_m \right\} ds$$

in the Euclidean case. The integrand is independent of the choice of the parameters u^a on C_ν.

In the spherical case, let the manifold C_ν be given by the parametric representation (6) with the normalization $\mathfrak{r}^2 = 1$. Therefore $\mathfrak{r} \cdot \mathfrak{r}_a = 0$. The mutually orthogonal normal vectors $\mathfrak{n} = \mathfrak{n}(1), \cdots, \mathfrak{n}(m)$ satisfy the equations

$$\mathfrak{r} \cdot \mathfrak{n} = 0, \qquad \mathfrak{r}_a \cdot \mathfrak{n} = 0.$$

From both equations there follows

$$\mathfrak{r} \cdot \mathfrak{n}_a = 0.$$

The part $C_\nu(\alpha)$ of the space S_n covered by the $m = (n - \nu)$-dimensional solid spheres of spherical radius α is represented by

$$\mathfrak{x} = \mathfrak{r} + t_1\mathfrak{n}(1) + \cdots + t_m\mathfrak{n}(m),$$

where the argument u in $\mathfrak{r}, \mathfrak{n}(1), \cdots, \mathfrak{n}(m)$ ranges over the whole C_ν, while the parameters $t_1, \cdots, t_m$ are bound by

$$t_1^2 + \cdots + t_m^2 \leqq a^2, \qquad (a = \tan\alpha).$$

According to equation (5) the volume $V(\alpha)$ of $C_\nu(\alpha)$ is given by the integral of

$$(9) \qquad \frac{[\mathfrak{x}\mathfrak{x}_1 \cdots \mathfrak{x}_\nu \mathfrak{n}(1) \cdots \mathfrak{n}(m)]}{(\mathfrak{x}^2)^{(n+1)/2}} du^1 \cdots du^\nu dt_1 \cdots dt_m$$

extended with respect to $u^1, \cdots, u^\nu$ over the whole of C_ν, with respect to $t_1, \cdots, t_m$ over the sphere $\sigma_m(a)$. Application of the same procedure as before results in the formula

$$(10) \quad V(\alpha) = \int_{C_\nu} \left\{ \int \cdots \int_{(t_1^2 + \cdots + t_m^2 \leqq a^2)} | \delta_\alpha{}^\beta + \sum_p t_p G_\alpha{}^\beta(p) | \times \frac{dt_1 \cdots dt_m}{(1 + t_1^2 + \cdots + t_m^2)^{(n+1)/2}} \right\} ds.$$

3. Evaluation. For any function $\phi(t) = \phi(t_1, \cdots, t_m)$, let $\langle\phi(t)\rangle_t$ designate its mean value over the sphere

$$(11) \qquad t_1^2 + \cdots + t_m^2 = 1.$$

The mean value $\langle t_1^{e_1} \cdots t_m^{e_m}\rangle_t$ of a monomial is obviously zero unless all exponents e_p are even. In the latter case one has the well-known formula

$$(12) \qquad \langle t_1^{e_1} \cdots t_m^{e_m}\rangle_t = \frac{e_1) \cdots e_m)}{m(m+2) \cdots (m+e-2)} \qquad (e_p \text{ even}, \ e = e_1 + \cdots + e_m),$$

where

$$0) = 1, \quad e) = 1 \cdot 3 \cdots (e-1) \qquad [\text{for } e = 2, 4, \cdots].$$

[(12) is most easily proved by multiplying the monomial by

$$e^{-t_1^2} \cdots e^{-t_m^2} = e^{-(t_1^2 + \cdots + t_m^2)}$$

and then integrating over

$$-\infty < t_p < \infty \qquad (p = 1, \cdots, m).$$

One thus obtains

$$\int t_1^{e_1} \cdots t_m^{e_m} d\omega_t \cdot \int_0^\infty e^{-r^2} r^{e+m-1} dr = \prod_p \left(\int_{-\infty}^{+\infty} e^{-t^2} t^{e_p} dt \right), \qquad (e = e_1 + \cdots + e_m),$$

with $\int \cdots d\omega_t$ indicating the "solid angle" integration over the sphere (11), and hence

$$\frac{1}{2} \int t_1^{e_1} \cdots t_m^{e_m} d\omega_t = \frac{\Gamma\left(\frac{1+e_1}{2}\right) \cdots \Gamma\left(\frac{1+e_m}{2}\right)}{\Gamma\left(\frac{m+e}{2}\right)}.$$

In particular, for the surface $\omega_m = \int d\omega_t$ of the sphere,

$$(13) \qquad \frac{1}{2} \omega_m = \left[\Gamma\left(\frac{1}{2}\right)\right]^m \Big/ \Gamma\left(\frac{m}{2}\right).$$

Division results in the desired equation

$$\langle t_1^{e_1} \cdots t_m^{e_m} \rangle_t = \frac{\Gamma\left(\frac{1}{2} + \frac{e_1}{2}\right) \cdots \Gamma\left(\frac{1}{2} + \frac{e_m}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \cdots \Gamma\left(\frac{1}{2}\right)} \Bigg/ \frac{\Gamma\left(\frac{m}{2} + \frac{e}{2}\right)}{\Gamma\left(\frac{m}{2}\right)}$$

$$= \frac{e_1) \cdots e_m)}{m(m+2) \cdots (m+e-2)}.$$

The volume of the solid sphere $\sigma_m(a)$ amounts to

$$\omega_m \int_0^a r^{m-1} dr = \frac{\omega_m}{m} \cdot a^m;$$

hence $\Omega_m = \omega_m / m$. Specialization of (13) for $m = 2$ yields $[\Gamma(\frac{1}{2})]^2 = \pi$. The numbers ω_m, Ω_m are best defined by the recursive formulas readily derived from (13):

$$\omega_{m+2} = \frac{2\pi}{m} \cdot \omega_m \quad (m \geqq 1); \qquad (\omega_1 = 2, \quad \omega_2 = 2\pi).$$

$$\Omega_{m+2} = \frac{2\pi}{m+2} \cdot \Omega_m \quad (m \geqq 0); \qquad (\Omega_0 = 1, \ \Omega_1 = 2).]$$

We expand the determinant

$$\psi(t_1 \cdots t_m) = | \delta_\alpha{}^\beta + \sum_p t_p G_\alpha{}^\beta(p) | = \psi_0 + \psi_1 + \cdots + \psi_\nu$$

according to degrees in the variables $t_1, \cdots, t_m$:

$$\psi_e(t_1 \cdots t_m) = \sum \phi_{e_1 \ldots e_m} t_1^{e_1} \cdots t_m^{e_m} \qquad (e_1 + \cdots + e_m = e)$$

is homogeneous of degree e. $\psi_0 = 1$. This decomposition is conveniently described by introducing an artificial parameter λ:

$$|\,\delta_a{}^\beta + \lambda \sum_p t_p G_a{}^\beta(p)\,| = 1 + \lambda\psi_1 + \lambda^2\psi_2 + \cdots.$$

We set

$$\langle \psi_e(t_1 \cdots t_m)\rangle_t = \frac{H_e}{m(m+2)\cdots(m+e-2)}.$$

By its definition, H_e is a point invariant of C_ν. H_e is zero for odd e, while for even e one derives from (12) the explicit expression

$$H_e = \sum e_1)\cdots e_m)\cdot \phi_{e_1 \ldots e_m}, \qquad (e_p \text{ even},\ e_1 + \cdots + e_m = e).$$

The integral over the solid sphere $\sigma_m(a)$,

$$\int \cdots \int_{\sigma_m(a)} \psi_e(t_1 \cdots t_m)\, dt_1 \cdots dt_m$$

then will turn out to be

$$\frac{\omega_m H_e}{m(m+2)\cdots(m+e-2)} \int_0^a r^{e+m-1} dr = \omega_m H_e \cdot \frac{a^{m+e}}{m(m+2)\cdots(m+e)}.$$

Thus we find in the Euclidean case

$$(14) \qquad V(a) = \Omega_m \sum_e k_e \frac{a^{m+e}}{(m+2)(m+4)\cdots(m+e)}, \qquad (e \text{ even},\ 0 \leqq e \leqq \nu),$$

with the coefficients

$$(15) \qquad k_e = \int_{C_\nu} H_e ds.$$

In the spherical case one gets

$$\int \cdots \int_{\sigma_m(a)} \psi_e(t) \frac{dt_1 \cdots dt_m}{(1 + t_1^2 + \cdots + t_m^2)^{(n+1/)2}} = \frac{\omega_m H_e}{m(m+2)\cdots(m+e-2)} \int_0^a \frac{r^{e+m-1} dr}{(1+r^2)^{(n+1/)2}}.$$

On putting $r = \tan\rho$ the integral at the right side becomes

$$\int_0^a (\sin\rho)^{m+e-1} (\cos\rho)^{\nu-e} d\rho,$$

and instead of (14) one obtains

$$(16) \qquad V(\alpha) = \omega_m \cdot \sum_e k_e J_e(\alpha), \qquad (e \text{ even},\ 0 \leqq e \leqq \nu),$$

where

$$m(m+2)\cdots(m+e-2)J_e(\alpha)=\int_0^\alpha(\sin\rho)^{m+e-1}(\cos\rho)^{\nu-e}d\rho. \tag{17}$$

One may notice the recurrent equation

$$\frac{(\sin\alpha)^{e+m}(\cos\alpha)^{\nu-e-1}}{m(m+2)\cdots(m+e)}=J_e(\alpha)-(\nu-e-1)J_{e+2}(\alpha).$$

THEOREM. *The volumes of (n,ν)-tubes in Euclidean and in spherical space are given by the formulas* (14), (16) *respectively,* $J_e(\alpha)$ *being defined by* (17). k_e, (15), *are certain integral invariants of* C_ν, *in particular* k_0 *is its surface.*

4. Intrinsic nature of the invariants k_e. So far we have hardly done more than what could have been accomplished by any student in a course of calculus. However, some less obvious argument is needed for ascertaining that more explicit form of the point invariant H_e which enables one to replace the curvature $G_\alpha{}^\beta(p)$ by the Riemannian tensor $R^\lambda{}_{\mu\alpha\beta}$ of C_ν. I repeat the definition of this tensor in terms of the metric ground tensor:

$$\sum_\lambda g_{\kappa\lambda}\Gamma^\lambda{}_{\alpha\beta}=\tfrac{1}{2}\left(\frac{\partial g_{\alpha\kappa}}{\partial u^\beta}+\frac{\partial g_{\beta\kappa}}{\partial u^\alpha}-\frac{\partial g_{\alpha\beta}}{\partial u^\kappa}\right)$$

[definition of the affine connection $\Gamma^\kappa{}_{\alpha\beta}$],

$$R^\kappa{}_{\lambda\alpha\beta}=\left(\frac{\partial\Gamma^\kappa{}_{\lambda\beta}}{\partial u^\alpha}-\frac{\partial\Gamma^\kappa{}_{\lambda\alpha}}{\partial u^\beta}\right)+\sum_\rho(\Gamma^\kappa{}_{\rho\alpha}\Gamma^\rho{}_{\lambda\beta}-\Gamma^\kappa{}_{\rho\beta}\Gamma^\rho{}_{\lambda\alpha}).$$

After raising the index λ according to

$$R^\kappa{}_{\lambda\alpha\beta}=\sum_\mu g_{\lambda\mu}R^{\kappa\mu}_{\alpha\beta},$$

$R^{\kappa\lambda}_{\alpha\beta}$ is not only skew-symmetric in $\alpha\beta$, but also in $\kappa\lambda$. As a part of the integrability conditions expressing the *Euclidean* nature of the embedding space E_n, one has the relations [2]

$$R^{\kappa\lambda}_{\alpha\beta}=\sum_{p=1}^m\{G_\alpha{}^\kappa(p)G_\beta{}^\lambda(p)-G_\beta{}^\kappa(p)G_\alpha{}^\lambda(p)\}. \tag{18}$$

In the *spherical* case we look upon C_ν as a surface in E_{n+1}. To the set of m normals $\mathfrak{n}(p)$, $(p=1,\cdots,m)$ one has simply to add $\mathfrak{n}(0)=\mathfrak{r}$. Since

[2] See H. Weyl, *Mathematische Zeitschrift*, vol. 12 (1922), p. 154.

$$\mathfrak{n}_\alpha(0) = \mathfrak{r}_\alpha \text{ or } G_\alpha{}^\beta(0) = \delta_\alpha{}^\beta,$$

(18) changes into the equation

$$(19) \quad R^{\kappa\lambda}_{\alpha\beta} - (\delta_\alpha{}^\kappa\delta_\beta{}^\lambda - \delta_\beta{}^\kappa\delta_\alpha{}^\lambda) = \sum_{p=1}^{m} \{G_\alpha{}^\kappa(p)G_\beta{}^\lambda(p) - G_\beta{}^\kappa(p)G_\alpha{}^\lambda(p)\}.$$

[It is a pity that the inadequate name "curvature," which ought to be reserved for $G_\alpha{}^\beta(p)$, has been attached to the Riemann tensor. In the paper just quoted I proposed the more descriptive term "vector vortex." The left side of (19), and also of (18), is the excess of the vortex of C_ν over that of the embedding space. In this form the relation would hold with an arbitrary embedding Riemann space.]

We must try then to express the spherical average

$$\langle \det (\delta_\alpha{}^\beta + \lambda \sum_p t_p G_\alpha{}^\beta(p))\rangle,$$

in terms of the quantities

$$(20) \qquad H\binom{\kappa\lambda}{\alpha\beta} = \sum_p G_\alpha{}^\kappa(p)G_\beta{}^\lambda(p) - \sum_p G_\alpha{}^\lambda(p)G_\beta{}^\kappa(p).$$

In this investigation the

$$G_\alpha{}^\beta = (G_\alpha{}^\beta(1), \cdots, G_\alpha{}^\beta(m)),$$

just as

$$t = (t_1, \cdots, t_m),$$

may be looked upon as arbitrary vectors in an m-dimensional Euclidean space E_m. Using for a moment the abbreviation

$$z_\alpha{}^\beta = (t \cdot G_\alpha{}^\beta) = \sum_p t_p G_\alpha{}^\beta(p),$$

one has

$$\psi_e = \sum_{a_1 < \cdots < a_e} \begin{vmatrix} z_{a_1}{}^{a_1}, & \cdots, & z_{a_1}{}^{a_e} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ z_{a_e}{}^{a_1}, & \cdots, & z_{a_e}{}^{a_e} \end{vmatrix}.$$

Hence we try to determine

$$\langle \det (t \cdot G_\alpha{}^\beta)\rangle_t,$$

where $G_\alpha{}^\beta$, $(\alpha, \beta = 1, \cdots, e)$ are any e^2 given vectors in E_m.

LEMMA.

(21) $\langle \det_{(\alpha,\beta=1,\cdots,e)} (t \cdot G_\alpha{}^\beta) \rangle,$

$$= \frac{1}{m(m+2)\cdots(m+e-2)} \sum_{[\alpha,\beta]} \delta\binom{\beta}{\alpha} H\binom{\beta_1\beta_2}{\alpha_1\alpha_2}\cdots H\binom{\beta_{e-1}\beta_e}{\alpha_{e-1}\alpha_e}.$$

$\alpha_1\cdots\alpha_e$, $\beta_1\cdots\beta_e$ are the numbers $1,\cdots,e$ in any two arrangements, $\delta\binom{\beta}{\alpha} = \pm 1$ according as the permutation carrying the α- into the β-arrangement is even or odd. The sum extends over all couplings of pairs

$$(22) \qquad \begin{array}{c|c|c} (\alpha_1\alpha_2) & (\alpha_3\alpha_4) & \\ (\beta_1\beta_2) & (\beta_3\beta_4) & \end{array} \cdots.$$

By a "pair" $(\alpha_1\alpha_2)$ we mean here two distinct numbers α_1, α_2, *irrespective* of their order. Indeed the term $T\binom{\beta}{\alpha}$ under the sum $\sum_{[\alpha\beta]}$ on the right side of (21) does not change under reversal of an α-pair, $(\alpha_1\alpha_2) \rightarrow (\alpha_2\alpha_1)$, or of a β-pair. Nor does it change under permutation of its $e/2$ factors H; therefore only the coupling of the α-pairs with the β-pairs, but not the order of the $e/2$ blocks of the scheme (22) matters. Of the $2^e \cdot (\frac{1}{2}e)\,!$ equal terms arising from $T\binom{\beta}{\alpha}$ by inverting any of the e pairs of indices and by permuting the $e/2$ factors H, only one is retained in the sum.

Taking the lemma for granted, we find at once

$$(23) \qquad H_e = \sum_{[\alpha,\beta]} \delta\binom{\beta}{\alpha} H\binom{\beta_1\beta_2}{\alpha_1\alpha_2}\cdots H\binom{\beta_{e-1}\beta_e}{\alpha_{e-1}\alpha_e},$$

where the sum now extends to all couplings of pairs (22) from the larger range $1, 2, \cdots, \nu$ for which the β-sequence consists of the same e distinct figures as the α-sequence. The invariant nature of the sum to the right is evidenced when we first write it as

$$\frac{1}{2^e(e/2)\,!} \sum_{\alpha_1,\ldots,\alpha_e} \sum_{1',\ldots,e'} \pm H\binom{\alpha_{1'}\,\alpha_{2'}}{\alpha_1\;\alpha_2} H\binom{\alpha_{3'}\,\alpha_{4'}}{\alpha_3\;\alpha_4}\cdots, \qquad (e/2 \text{ factors}),$$

the inner sum alternatingly running over the permutations $1', \cdots, e'$ of $1, \cdots, e$. The limitation of distinctness imposed upon $\alpha_1, \cdots, \alpha_e$ can be canceled, as the inner sum vanishes if two of the α's coincide. Hence

$$(24) \qquad H_e = \frac{1}{2^e(e/2)\,!} \sum_{1',\ldots,e'} \left\{ \pm \sum_{\alpha_1,\ldots,\alpha_e} H\binom{\alpha_{1'}\,\alpha_{2'}}{\alpha_1\;\alpha_2} H\binom{\alpha_{3'}\,\alpha_{4'}}{\alpha_3\;\alpha_4}\cdots \right\}.$$

The inner sum in which each α runs independently from 1 to ν is a scalar. We have thus arrived at the decisive

THEOREM. *The scalar* H_e *on* C_ν *is determined by the formulas* (23), (24) *where* $H\binom{\lambda\mu}{\alpha\beta}$ *is the Riemann tensor or vortex* $R^{\lambda\mu}_{\alpha\beta}$ *in the Euclidean case, and the vortex excess* (19) *in the spherical case.*

These metric scalars H_e deserve attention on their own merits: they are probably the simplest and most fundamental scalars built up by the Riemann tensor.

As a very special case of our theorem we find that *the one term formulas*

$$V(a) = \Omega_m a^m \cdot k_0, \qquad V(\alpha) = \omega_m J_0(\alpha) \cdot k_0$$

prevail if C_ν *is applicable on* E_ν *or* S_ν *respectively.* k_0 denotes the surface of C_ν. Professor Hotelling's result concerning the tubes around a *curve*, $\nu = 1$, is fully contained in this special case.

The lemma is proved by an invariant-theoretic argument as follows. We consider the e^2 vectors $G_\alpha{}^\beta$ as independent variables.

$$\Phi = \langle \det (t \cdot G_\alpha{}^\beta) \rangle_t$$

is an orthogonal invariant of these variables and therefore, according to the theory of orthogonal vector invariants,[3] expressible as a polynomial in terms of the scalar products $(G_\alpha{}^\lambda \cdot G_\beta{}^\mu)$. Observing that Φ is linear and homogeneous in the components of the vectors of each row and each column of the scheme

$$\left\| \begin{matrix} G_1{}^1, \cdots, G_1{}^e \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ G_e{}^1, \cdots, G_e{}^e \end{matrix} \right\|$$

we realize that it must be a linear combination of terms

$$(G_{\alpha_1}{}^{\beta_1} \cdot G_{\alpha_2}{}^{\beta_2}) \cdots (G^{\beta_{e-1}}_{\alpha_{e-1}} \cdot G_{\alpha_e}{}^{\beta_e}),$$

where the α and β are any two arrangements of $1, \cdots, e$. Moreover Φ is skew-symmetric with respect to the columns. Hence, by summing alternatingly over the $e!$ permutations of the superscripts β we find that Φ is a linear combination of the following functions

[3] E. Study, *Ber. Sächs. Akad. Wissensch.* 1897, p. 442. H. Weyl, *Mathematische Zeitschrift*, vol. 20 (1924), p. 136.

$$\sum_{(\beta)} \delta\binom{\beta}{\alpha} (G_{\alpha_1}{}^{\beta_1}\cdot G_{\alpha_2}{}^{\beta_2})(G_{\alpha_3}{}^{\beta_3}\cdot G_{\alpha_4}{}^{\beta_4})\cdots = \sum_{[\beta]} \delta\binom{\beta}{\alpha}\, H\binom{\beta_1\beta_2}{\alpha_1\alpha_2}\, H\binom{\beta_3\beta_4}{\alpha_3\alpha_4}\cdots.$$

The first sum runs over all $e!$ permutations $\beta_1\cdots\beta_e$ of $1,\cdots,e$, the second over all their $e!/2^{e/2}$ arrangements in "pairs"

$$(\beta_1\beta_2),\quad (\beta_3\beta_4),\cdots.$$

By applying the same argument to the subscripts α one concludes that Φ is a constant multiple c of H_e, (23).

The constant c is determined by the specialization

$$G_\alpha{}^\beta = (\delta_\alpha{}^\beta, 0,\cdots,0)$$

for which

$$\Phi = \langle t_1{}^e\rangle_t = \frac{e)}{m(m+2)\cdots(m+e-2)}$$

and

$$H\binom{\lambda\mu}{\alpha\beta} = \delta_\alpha{}^\lambda\delta_\beta{}^\mu - \delta_\alpha{}^\mu\delta_\beta{}^\lambda,\quad H_e = e!/2^{e/2}(\tfrac{1}{2}e)! = e).$$

Institute for Advanced Study,
Princeton, N. J.

NOTE ON POWER SERIES WITH BIG GAPS.*

By M. Kac.[1]

This note contains some results concerning the asymptotic distribution of the values of analytic functions $F(z)$, represented for $|z| < 1$ by the power series $F(z) = \Sigma a_k z^{n_k}$, where $n_k/(n_1 + \cdots + n_{k-1}) \to \infty$ as $k \to \infty$, and $\Sigma |a_k|^2 = \infty$. It will be proved that the distribution of values of $F(z)$ in the neighborhood of the convergence circle is, in a certain sense, the normal distribution. Thus, the connection between gap theorems and statistics will be made clear once more.[2]

In order to simplify the calculations one supposes that the a_k are real and that $a_k = O(1)$.

From the fact that $a_k = O(1)$ and $\Sigma a_k^2 = \infty$, it follows that the radius of convergence is 1.

1. Let $f_k(t)$ be independent functions[3] on the interval $[0, 1]$ and let

$$\int_0^1 f_k(t)\,dt = 0, \quad \int_0^1 f_k^2(t)\,dt = h^2, \quad |f_k(t)| < A, \quad (k = 1, 2, \cdots).$$

Putting $M(r) = (\sum_{k=1}^{\infty} a_k^2 r^{n_k})^{\frac{1}{2}}$, one has, for every integer $l \geqq 0$,

$$\lim_{r\to 1} M^{-l}(r) \int_0^1 (\sum_{k=1}^{\infty} a_k r^{n_k} f_k(t))^l\,dt = h^l 2^{l/2} \pi^{-1/2} \int_{-\infty}^{\infty} u^l e^{-u^2} du.$$

In order to prove this, notice first that the statistical independence of $f_k(t)$ and the condition $|f_k(t)| < A$ imply that, for every complex z,

* Presented to the American Mathematical Society February 25, 1939. Received January 12, 1939.

[1] Fellow of the Parnas Foundation, Lwów, Poland.

[2] R. E. A. C. Paley and A. Zygmund, "On some series of functions (1) and (2)," *Proceedings of the Cambridge Philosophical Society*, vol. 26 (1930), pp. 337-357 and 458-474. See also N. Wiener, "Gap theorems," *Comptes Rendus du Congrès International des Mathématiciens*, Oslo 1936.

[3] The measurable functions $f_1(t), \cdots, f_k(t)$ defined on $[0, 1]$ are called independent if for any system of real numbers $a_1, \cdots, a_k$ one has

$$|E\{f_1(t) < a_1, \cdots, f_k(t) < a_k\}| = \prod_{j=1}^{k} |E\{f_j(t) < a_j\}|,$$

where $E\{\ \}$ denotes the set of those t for which the relation inside of $\{\ \}$ is satisfied, and $|E\{\ \}|$ the measure of this set. For properties of independent functions, see for instance M. Kac, "Sur les fonctions indépendantes I," *Studia Mathematica*, vol. 6 (1936), pp. 46-58.

$$\int_0^1 \exp(zM^{-1}(r) \sum_{k=1}^{\infty} a_k r^{n_k} f_k(t))\,dt = \prod_{k=1}^{\infty} \int_0^1 \exp(zM^{-1}(r) a_k r^{n_k} f_k(t))\,dt.$$

Furthermore,

$$\int_0^1 \exp(zM^{-1}(r) a_k r^{n_k} f_k(t))\,dt = 1 + \tfrac{1}{2}h^2 z^2 M^{-2}(r) a_k^2 r^{2n_k} + z^3 o\,(M^{-2}(r) a_k^2 r^{2n_k}),$$

where the o-term is uniform in k, since $|f_k(t)| < A$. Hence it is readily seen that

$$\int_0^1 \exp(zM^{-1}(r) \sum_{k=1}^{\infty} a_k r^{n_k} f_k(t))\,dt \to \exp(\tfrac{1}{2}h^2 z^2) \text{ as } r \to 1,$$

holds uniformly in every finite circle $|z| \leq R$. Now, the coefficient of z^l in the expansion on the left tends to the corresponding coefficient on the right. This completes the proof.

2. Next, it will be proved that the relation, just proved, also holds if one replaces $f_k(t)$ by $\cos 2\pi n_k t$, even though the cosines are not independent.

Let $\vartheta_1(t), \vartheta_2(t), \cdots$ be the independent functions of Steinhaus,[4] which realize the mapping of $(0,1)$ on the infinite dimensional torus and conserve measure. Putting $f_k(t) = \cos 2\pi\vartheta_k(t)$, one obtains a sequence of independent functions satisfying all conditions of § 1 with $h^2 = 1/2$. Then

$$(*)\qquad \lim_{r\to 1} M^{-l}(r) \int_0^1 (\sum_{k=1}^{\infty} a_k r^{n_k} \cos 2\pi\vartheta_k(t))^l\,dt = \pi^{-1/2} \int_{-\infty}^{+\infty} u^l e^{-u^2}\,du.$$

For a fixed l, choose a $k(l)$ in such a way that $n_k/(n_1 + \cdots + n_{k-1}) > l$ for every $k > k(l)$. It will first be proved that

$$(**)\qquad \int_0^1 (\sum_{k=k(l)}^{\infty} a_k r^{n_k} \cos 2\pi n_k t)^l\,dt = \int_0^1 (\sum_{k=k(l)}^{\infty} a_k r^{n_k} \cos 2\pi\vartheta_k(t))^l\,dt.$$

To this end, it is sufficient to show that

$$\int_0^1 \cos^{p_1} 2\pi n_{k_1} t \cdots \cos^{p_l} 2\pi n_{k_l} t\,dt = \int_0^1 \cos^{p_1} 2\pi\vartheta_{k_1}(t) \cdots \cos^{p_l} 2\pi\vartheta_{k_l}(t)\,dt,$$

where $p_1 + p_2 + \cdots + p_l = l$, $k_s > k(l)$. Observing that

$$\delta(p_1, \cdots, p_l) \equiv \int_0^1 \cos^{p_1} 2\pi\vartheta_{k_1}(t) \cdots \cos^{p_l} 2\pi\vartheta_{k_l}(t)\,dt = \prod_{s=1}^{l} \int_0^1 \cos^{p_s} 2\pi t\,dt,$$

one evidently has $\delta(p_1, \cdots, p_l) = 0$, if at least one p_s is odd; while

$$\delta(p_1, \cdots, p_l) = 2^{-(p_1 + \cdots + p_l)} \binom{p_1}{p_1/2} \cdots \binom{p_l}{p_l/2},$$

[4] H. Steinhaus, "Sur la probabilité de la convergence de séries," *Studia Mathematica*, vol. 2 (1930), pp. 21-39.

if all p are even. Thus, the establishing of the statement (**) depends on showing that

$$\int_1^1 \cos^{p_1} 2\pi n_{k_1} t \cdots \cos^{p_l} 2\pi n_{k_l} t \, dt = \delta(p_1, \cdots, p_l).$$

Using the complex formula for cosines, one obtains

$$\cos^{p_1} 2\pi n_{k_1} t \cdots \cos^{p_l} 2\pi n_{k_l} t = 2^{-(p_1 + \cdots + p_l)} \sum \binom{p_1}{s_1} \cdots \binom{p_l}{s_l} e^{2\pi i \lambda t},$$

where the summation indices s_i run from 0 to p_i, and

$$\lambda = (2s_1 - p_1) n_{k_1} + \cdots + (2s_l - p_l) n_{k_l}.$$

Since $\int_0^1 \exp(2\pi i \lambda t) dt$ is 0 or 1 according as $\lambda \neq 0$ or $\lambda = 0$, it remains to be proved that $\lambda = 0$ if and only if $s_1 = p_1/2, \cdots, s_l = p_l/2$. This is obvious, since $n_k/(n_1 + \cdots + n_{k-1}) > l$ for $k > k(l)$ and $k_s > k(l)$. In fact, suppose that not all $2s_j - p_j$ vanish and let m be the largest index for which $2s_m - p_m \neq 0$, then

$$|\lambda| \geq \left| \, |2s_m - p_m| \, n_{k_m} - \sum_{j=1}^{m-1} |2s_j - p_j| \, n_{k_j} \right|,$$

$$\sum_{j=1}^{m-1} |2s_j - p_j| \, n_{k_j} \leq l \sum_{j=1}^{m-1} n_{k_j} < l \sum_{j=1}^{k_{m-1}} n_j < n_{k_m},$$

and, as $|2s_m - p_m| \, n_{k_m} \geq n_{k_m}$, one has $|\lambda| > 0$.

4. In view of the known inequality [5]

$$\int_0^1 |\sum c_k \cos 2\pi \vartheta_k(t)|^p < c(p)(\sum c_k^2)^{p/2},$$

one has

$$\int_0^1 \left| \sum_{k=k(l)}^{\infty} a_k r^{n_k} \cos 2\pi \vartheta_k(t) \right|^p dt < c(p) M^p(r).$$

An analogous inequality holds [6] for

$$\int_0^1 \left| \sum_{k=k(l)}^{\infty} a_k r^{n_k} \cos 2\pi n_k t \right|^p dt, \quad \text{if} \quad p < l.$$

Observing that, as $r \to 1$,

$$M^{-1}(r) \sum_{k=1}^{k(l)-1} a_k r^{n_k} \cos 2\pi \vartheta_k(t) \quad \text{and} \quad M^{-1}(r) \sum_{k=1}^{k(l)-1} a_k r^{n_k} \cos 2\pi n_k t$$

tend to 0 uniformly in t, one readily infers from (*) and (**) that

---

[5] *Loc. cit.* 1), pp. 467-468.

[6] This is an immediate consequence of (**) for even p. In the case p is odd, one can apply Hölder's inequality.

$$\lim_{r\to 1} M^{-l}(r)\int_0^1 \Big(\sum_{k=1}^{\infty} a_k r^{n_k}\cos 2\pi n_k t\Big)^l\, dt = \pi^{-1/2}\int_{-\infty}^{+\infty} u^l e^{-u^2} du.$$

This implies, according to a well known theorem,[7] that

$$\lim_{r\to 1} \Big| E\{\sum_{k=1}^{\infty} a_k r^{n_k}\cos 2\pi n_k t < \omega M(r)\}\Big| = \pi^{-1/2}\int_{-\infty}^{\omega} e^{-u^2} du.$$

Putting $2\pi t = \phi$ one has

$$\Re(F(re^{i\phi})) = \sum_{k=1}^{\infty} a_k r^{n_k}\cos n_k\phi = \sum_{k=1}^{\infty} a_k r^{n_k}\cos 2\pi n_k t$$

and one arrives at the following theorem:

$$\lim_{r\to 1} \big| E_\phi\{\Re(F(re^{i\phi})) < \omega M(r)\}\big| = 2\pi^{1/2}\int_{-\infty}^{\omega} e^{-u^2} du.$$

A similar reasoning also proves the same theorem for $\Im(F(re^{i\phi}))$.

Using the same method with small modifications, one readily finds the sharper relation

$$\lim_{r\to 1} \big| E\{\Re(F(re^{i\phi}) < \omega_1 M(r), \Im(F(re^{i\phi})) < \omega_2 M(r)\}\big| = 4\pi\int_{-\infty}^{\omega_1} e^{-u^2} du \int_{-\infty}^{\omega_2} e^{-u^2} du.$$

5. It is easy to see that if the gap condition is unaltered the above method permits one to obtain

$$\lim_{m\to\infty} m^{-l/2}\int_0^1 \Big(\sum_{k=1}^{m}\cos 2\pi n_k t\Big)^l dt = \pi^{-\frac{1}{2}}\int_{-\infty}^{+\infty} u^l e^{-u^2} du,$$

and therefore

$$\lim_{m\to\infty} \Big| E\{\sum_{k=1}^{m}\cos 2\pi n_k t < \omega m^{\frac{1}{2}}\}\Big| = \pi^{-1/2}\int_{-\infty}^{\omega} e^{-u^2} du.$$

It is clear that a similar theorem may be obtained for Dirichlet series with the gap condition considered above.

It seems to be not without interest to investigate similar questions in the case of the usual gap condition $n_k/n_{k-1} > q > 1$.

Added March 22, 1939.—Professor Wintner has pointed out to me that one can replace the gap considerations, used above, by the corresponding considerations of Paley and Zygmund (cf. A. Zygmund, *Fundamenta Mathematica*, vol. 16 (1930), p. 104 and top of p. 105) and that the result of the present paper may then be obtained without recourse to independent functions. In this way the gap condition $n_k/(n_1 + \cdots + n_{k-1}) \to \infty$, used above, becomes replaced by the condition $n_k/n_{k-1} \to \infty$, which seems to be more general. Actually it is easy to see that these conditions are equivalent.

THE JOHNS HOPKINS UNIVERSITY.

[7] Compare for instance A. A. Markoff's book on calculus of probability, and, in the multidimensional case, E. K. Haviland, *American Journal of Mathematics*, vol. 56 (1935), pp. 625-658.

ASYMPTOTIC DISTRIBUTIONS AND STATISTICAL INDEPENDENCE.*

By Philip Hartman, E. R. van Kampen and Aurel Wintner.

Introduction. The present paper deals with the distribution problem of functions which are independent in the sense of Kolmogoroff [1] or, equivalently, in that of Steinhaus.[2] For such functions, Steinhaus has developed, in coöperation with Kac, a valuable theory [2] in case the functions are defined on a finite interval.[3] However, the case which is significant in physical and number theoretical applications is the case of functions on the infinite range $-\infty < t < +\infty$; so that the integration processes must be replaced by averaging processes. In this case, the theory developed [4] leads to certain difficulties, which will be analyzed in § 9 (and § 2 bis) below.

The purpose of the present paper is to remove these difficulties, thus making the theory of independent functions applicable to the fields mentioned above. This aim will be reached by establishing the connection between the theory of independent functions on the one hand and the theory of convolutions [5] on the other hand (cf. the theorem of § 7 below).

In § 9, it will be necessary to refer to the simplest case of a function with an asymptotic distribution function, viz., the case of a real-valued uniformly almost periodic function $f(t)$, $-\infty < t < +\infty$. If $\sigma_T(x)$, where $T > 0$ and $-\infty < x < +\infty$, denotes the ratio of the sum of the lengths of those subintervals of the interval $-T \leqq t \leqq T$ on which $f(t) < x$ and of the length $2T$ of this interval, then $f(t)$ has [6] an asymptotic distribution function, in the sense that there exists a monotone function $\sigma(x)$, $-\infty < x < +\infty$, such that $\sigma_T(x)$ tends, as $T \to \infty$, to $\sigma(x)$ with the possible exception of those

* Received February 11, 1939.

[1] A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin (1933), p. 50.

[2] Cf. H. Steinhaus, *Actualités Scientifiques et Industrielles* (1938), where further references are given.

[3] M. Kac, *Studia Mathematica*, vol. 6 (1936), pp. 46-58.

[4] M. Kac and H. Steinhaus, *Studia Mathematica*, vol. 7 (1937), pp. 1-15.

[5] Cf., e. g., B. Jessen and A. Wintner, *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88 (more particularly § 2-§ 3), where further references are given.

[6] A. Wintner, *Zeitschrift für Physik*, vol. 48 (1928), pp. 148-161; *Mathematische Zeitschrift*, vol. 30 (1929), pp. 290-319.

points x which are discontinuity points of $\sigma(x)$. (In particular, the set of those x for which $\lim \sigma_T(x)$ does not exist is at most enumerable). It is clear from the example of a periodic, continuous, Cantor function that $\lim \sigma_T(x)$ may exist for every x even if the discontinuity points of this limit function $\sigma(x)$ are dense between $x = \operatorname{Min} f(t)$ and $x = \operatorname{Max} f(t)$. On the other hand, Bohr [7] has constructed an almost periodic function, which actually is limit-periodic (grenzperiodisch), with the property that the limit of $\sigma_T(x)$ does not exist at an exceptional x. (It is an unsolved problem whether or not such exceptional x may form, for a suitable almost periodic $f(t)$, a set which is dense on an interval). Obviously, one should like to consider the functions $f(t)$ and $f(rt)$ as statistically independent if r is an irrational number and $f(t)$ a limit-periodic function.

The theory to be developed takes care of the difficulties implied by the above situation.

1. Let R_x denote a real k-dimensional vector space of the points $x = (x^1, \cdots, x^k)$. By $x_I < x_{II}$ will be meant that $x_I{}^j < x_{II}{}^j$ for $j = 1, \cdots, k$. Correspondingly, $Q(a, b)$ will denote the k-dimensional interval $a^j < x^j < b^j$, $j = 1, \cdots, k$. A sequence $\{Q_n\}$ of intervals $Q_n = Q(a_n, b_n)$ will be said to be dense on R_x, if there exists for every interval $Q(\alpha, \beta) \subset R_x$ and for every $\epsilon > 0$ an $n = n(\alpha, \beta)$ such that $| \alpha - a_n | < \epsilon$ and $| \beta - b_n | < \epsilon$, where $| c |$ denotes the length of the vector c.

By a distribution function on R_x will be meant an additive monotone set function $\phi = \phi(Q)$, defined on a dense set of intervals Q in such a way that $\phi(R_x) = 1$. It is well known [8] that ϕ can uniquely be extended to all intervals Q which are "continuity intervals" of ϕ; and that "nearly all" intervals Q are continuity intervals of ϕ, in the sense that there exists for every ϕ a sequence of vectors, $\{\gamma_n\}$, in such a way that the interval $Q(a, b)$ is a continuity interval of ϕ whenever $a^j \neq \gamma_n{}^j \neq b^j$ for $j = 1, \cdots, k$ and $n = 1, 2, \cdots$. Correspondingly, a family $\phi_T(Q)$, $0 < T < \infty$, of distribution functions will be said to tend to a limit distribution function ϕ, if $\phi_T(Q) \rightarrow \phi(Q)$, where $T \rightarrow \infty$, holds at every continuity interval Q of ϕ.

In view of the continuity theorem of Fourier-Stieltjes transforms,[9] the limit distribution function ϕ exists if and only if the Fourier-Stieltjes transform $L(u; \phi_T)$ tends, as $T \rightarrow \infty$, to a limit function uniformly in every fixed sphere $| u | <$ const., in which case $\lim L(u; \phi_T) = L(u; \phi)$. It is under-

[7] H. Bohr, *Danske Videnskabernes Selskab, Mathematiske-Fysiske Meddelelser*, vol. 10, Nr. 6 (1930), pp. 12-17.

[8] Cf., e. g., A. Kolmogoroff, *loc. cit.* [1], Chap. II.

[9] Cf. E. K. Haviland, *American Journal of Mathematics*, vol. 57 (1935), pp. 382-388.

stood that $L(u;\psi)$ is defined, for every distribution function ψ and for every point $u=(u^1,\cdots,u^k)$ of a real k-dimensional vector space R_u, by the k-dimensional Stieltjes integral

$$L(u;\psi)=\int_{R_x}\exp(iu\cdot x)\psi(dR_x), \tag{1}$$

where the dot denotes scalar multiplication. The value of ψ for a continuity interval Q of ψ is represented by the inversion formula [10] for the transform (1).

2. For a given measurable set E on a t-axis and for any $T>0$, let $\mu_T E$ denote the ratio of the Lebesgue measure of the common part of E and the interval $-T\leqq t\leqq T$ and of the length $2T$ of this interval; so that $0\leqq\mu_T E\leqq 1$. If $\mu_T E$ tends to a limit as $T\to\infty$, the measurable set E will be called relatively measurable,[11] and $\lim\mu_T E$ will be denoted by μE. Trivial examples show that the relative measure μ, though additive, is not completely additive, and that the common part of two relatively measurable sets need not be relatively measurable.

Let $x(t)$, $-\infty<t<+\infty$, be a given measurable vector function whose components $x^j(t)$ are coördinates x^j of a point $x\subset R_x$. For any $Q=Q(a,b)$, let $[x(t)\subset Q]_T$, where $T>0$, denote the set of those points t of the interval $-T\leqq t\leqq T$ at which $a<x(t)<b$. Since this set is measurable, one can define a family $\phi_T(Q)$, $0<T<+\infty$, of distribution functions by placing

$$\phi_T(Q)=\mu_T[x(t)\subset Q]_T. \tag{2}$$

The notation (1), when applied to (2), is equivalent to

$$L(u;\phi_T)=M_T\{\exp iu\cdot x(t)\}, \tag{3}$$

where $M_T\{g(t)\}=(1/2T)\int_{-T}^{T}g(t)\,dt.$

Now, if there exists a distribution function ϕ such that $\phi_T\to\phi$ as $T\to\infty$, then $x(t)$ is said to have an asymptotic distribution function, ϕ. In view of (3) and the continuity theorem of Fourier-Stieltjes transforms, this will be the case [11] if and only if the limit $M\{\exp iu\cdot x(t)\}$, where $M=\lim M_T$, exists uniformly in every fixed sphere $|u|<$ const., in which case

$$L(u;\phi)=M\{\exp iu\cdot x(t)\}. \tag{4}$$

And ϕ may be obtained from (4) by an application of the inversion formula of the transform (1).

---

[10] Cf. E. K. Haviland, *ibid.*, pp. 94-100.

[11] Cf. *loc. cit.* [5], pp. 74-75.

2 bis. Notice that (2) may tend, as $T \to \infty$, to a monotone additive set function also when $x(t)$ does not have an asymptotic distribution function. In fact, examples of the type $x(t) = t$ (or $x(t) = t \sin t$) show that the limit ($\equiv 0$) of (2) need not be a distribution function. (And the continuity theorem of the Fourier-Stieltjes transforms has no analogue in this case, as shown by trivial examples).

Thus, on using the notion of relative measure, and denoting by $[x(t) \subset Q]$, where $Q = Q(a, b)$, the set of those points of the t-axis at which $a < x(t) < b$, one can say that the relative measurability of the set $[x(t) \subset Q]$ for a dense set of intervals Q of R_x is necessary but not sufficient for the existence of an asymptotic distribution function ϕ of $x(t)$. In fact, one must also require that the total variation of

$$(5) \qquad \phi(Q) = \lim_{T\to\infty} \mu_T[x(t) \subset Q]_T = \mu[x(t) \subset Q]$$

over R_x (i. e., the total probability) should be 1.

3. For a given interval $Q = Q(a, b)$ of the k-dimensional space R_x, let Q^j, where $j = 1, \cdots, k$, denote the interval $a^j < x^j < b^j$ on the x^j-axis, R_{x^j}; so that $Q = Q^1 \times \cdots \times Q^k$. Let $x(t)$ be a measurable vector function which has an asymptotic distribution function. Then each of the k components $x^j(t)$ of $x(t)$ has an asymptotic distribution function

$$(6) \qquad \phi^j(Q^j) = \lim_{T\to\infty} \mu_T[x^j(t) \subset Q^j]_T = \mu[x^j(t) \subset Q^j], \qquad (j = 1, \cdots, k),$$

(where it is understood that the limits (5) and (6) only need to exist on "nearly all" intervals Q, Q^j of R_x, R_{x^j} respectively). This is clear from the M-criterion of § 2, since the uniform existence of the limit $M\{\exp iu \cdot x(t)\}$ on every fixed sphere $| u | <$ const., where $u = (u^1, \cdots, u^k)$, implies the uniform existence of the limit $M\{\exp iu^j x^j(t)\}$ on every fixed interval $| u^j | <$ const. (In fact, put in the scalar product $u \cdot x(t)$ all but the j-th component of u equal to 0).

On the other hand, it is quite possible that each of the k components $x^j(t)$ of $x(t)$ has, while $x(t)$ itself does not have, an asymptotic distribution function. In order to see this, it is sufficient to choose on the t-axis k sets $S^1, S^2, \cdots, S^k$ in such a way that while each of them is, their common part is not, relatively measurable, and then define $x^j(t)$ to be the characteristic function of S^j, where $j = 1, \cdots, k$.

4. Let k real, measurable, scalar functions $x^j(t)$, $-\infty < t < +\infty$, be called statistically independent, if, on the one hand, the vector function $x(t) = (x^1(t), \cdots, x^k(t))$ has an asymptotic distribution function (5), and,

on the other hand, (5) may be expressed in terms of the k asymptotic distribution functions (6) as follows:

$$(7) \qquad \phi(Q) \equiv \phi(Q^1 \times Q^2 \times \cdots \times Q^k) = \phi^1(Q^1)\phi^2(Q^2) \cdots \phi^k(Q^k).$$

It is clear that if the k functions $x^j(t)$ are statistically independent, then so are any $k-1$ functions in this set of k functions. On the other hand, if each of the k sets consisting of $k-1$ of k functions $x^j(t)$ represent statistically independent functions, then the k functions $x^j(t)$ need not be statistically independent. This is shown[14] by the example of the $k=3$ binary Walsh products

$$x^1(t) = r(t)r(2t), \qquad x^2(t) = r(2t)r(4t), \qquad x^3(t) = r(4t)r(t)$$

of Rademacher functions $r(2^n t)$, where $r(t) = \operatorname{sgn} \sin 2\pi t$; and, correspondingly, by the continuous example of the $k=3$ periodic functions

$$x^1(t) = \cos t, \qquad x^2(t) = \cos \pi t, \qquad x^3(t) = \cos(1+\pi)t.$$

5. Let k real, measurable, scalar functions on $-\infty < t < +\infty$ be called statistically independent in the additive sense, if, on the one hand, each of them has an asymptotic distribution function, and, on the other hand, the sum of them has an asymptotic distribution function which is represented by the convolution of the k one-dimensional distribution functions.

In this definition, the convolution $\psi = \psi^1 * \cdots * \psi^k$ of k one-dimensional distribution functions is meant in the usual sense[5] and denotes, therefore, the distribution function which may be defined by

$$(8) \qquad L(\lambda;\psi) = \prod_{j=1}^{k} L(\lambda;\psi^j), \text{ where } -\infty < \lambda < +\infty; \text{ cf. (1).}$$

In § 7, it will be convenient to think of the convolution of k one-dimensional distribution functions in a geometrical fashion, as follows:

If Q^j is an interval on the line R_{x^j} and $\psi^j = \psi^j(Q^j)$ a distribution function on R_{x^j}, where $j = 1, \cdots, k$, let $\sigma^j = \sigma^j(x^j)$, where $-\infty < x^j < +\infty$, denote the point function which represents the value of the set function ψ^j for the half-line whose upper end is the point x^j. For a given real number u^j, let $\sigma_{u^j}{}^j = \sigma_{u^j}{}^j(x^j)$ denote the point function on R_{x^j} which is equal to $\sigma^j(x^j/u^j)$, $1-\sigma^j(x^j/u^j)$ or $\frac{1}{2}(1+\operatorname{sgn} x^j)$ according as $u^j > 0$, $u^j < 0$ or $u^j = 0$. Then the one-dimensional distribution function, say ρ, which is the convolution of the k distribution functions $\psi_{u^j}{}^j$ associated with the k point func-

[14] Cf. S. Bernstein's example, mentioned *loc. cit.* [1], p. 10.

tions $\sigma_{u^j}{}^{j}$ may be obtained as follows: Consider on $R_x = R_{x^1} \times R_{x^2} \times \cdots \times R_{x^k}$ the distribution function $\psi(Q)$ which is defined by

$$\psi(Q^1 \times Q^2 \times \cdots \times Q^k) = \psi^1(Q^1)\psi^2(Q^2) \cdots \psi^k(Q^k),$$

and project this distribution in R_x orthogonally on that line, Γ_u, through the origin of R_x whose direction cosines are proportional to the components u^j of u, where $u \neq (0, \cdots, 0)$. Then the resulting one-dimensional distribution on Γ_u is precisely ρ.

This may easily be verified either directly [12] or by using Fourier-Stieltjes transforms.[13]

6. The notions of "statistical independence" (§ 4) and "statistical independence in the additive sense" (§ 5) are not equivalent. In fact, let [15]

$$x^1(t) = 0,\ \tfrac{1}{4},\ \tfrac{1}{2} \text{ for } 0 \leqq t < \tfrac{1}{3},\ \tfrac{1}{3} \leqq t < \tfrac{2}{3},\ \tfrac{2}{3} \leqq t < 1$$

and

$$x^2(t) = -\tfrac{1}{4},\ 0,\ -\tfrac{1}{2} \text{ for } \tfrac{1}{9} \leqq t < \tfrac{4}{9},\ \tfrac{4}{9} \leqq t < \tfrac{7}{9},\ \tfrac{7}{9} \leqq t < \tfrac{10}{9}$$

respectively, while

$$x^1(t) = x^1(t+1) \text{ and } x^2(t) = x^2(t+1) \text{ for } -\infty < t < +\infty.$$

It is easy to see that these $k = 2$ periodic step functions $x^1(t)$, $x^2(t)$ are statistically independent in the additive sense, without being statistically independent. It may be mentioned that for this $x(t) = (x^1(t), x^2(t))$ the functions $u^1x^1(t)$, $u^2x^2(t)$, where u^1, u^2 denote arbitrary distinct non-vanishing constants, are not independent in either sense.

---

[12] An early instance of this interpretation of the convolution process is implied by Sommerfeld's approach to the Gaussian law in the particular case of the "Abrundungsfehler" (*Boltzmann Festschrift*, Leipzig (1904), pp. 848-859 or *Bulletin of the Calcutta Mathematical Society*, vol. 20 (1930); cf. also V. Brun, *Norsk Matematisk Tidsskrift*, vol. 14 (1932), pp. 81-92 and F. Tricomi, Jahresbericht der Deutschen Mathematiker Vereinigung, vol. 42 (1933), pp. 174-179).

[13] Cf., e. g., H. Cramér and H. Woldt, *Journal of the London Mathematical Society*, vol. 11 (1936), pp. 290-294.

[15] Another example, based on $k = 3$ *infinite* Rademacher sums, was communicated to us by Dr. M. Kac; they correspond to the $k = 3$ Besicovitch almost periodic functions

$$x^j(t) \sim \sum_{n=1}^{\infty} a_n{}^j \cos \lambda_n t; \quad j = 1, 2, 3, \qquad \left(\sum_{n=1}^{\infty} (a_n{}^j)^2 < +\infty\right),$$

in which $\{\lambda_n\}$ is any sequence of real numbers which are linearly independent in the rational field (e. g., $\lambda_n = \pi^n$), and the three sequences $\{a_n{}^j\}$ are defined as follows:

$$\{a_n{}^1\}: \quad 1;\ \tfrac{1}{2}, \tfrac{1}{2};\ \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4};\ \tfrac{1}{8}, \cdots;\ \tfrac{1}{16}, \cdots$$

while $a_n{}^2 = a_n{}^1$ or $a_n{}^2 = -a_n{}^1$ according as $n > 1$ or $n = 1$, finally $a_n{}^3 = a_n{}^1 + a_n{}^2$ for every n.

7. We shall, however, prove the following theorem:

k real, measurable, scalar functions $x^j(t)$, $-\infty < t < +\infty$, are statistically independent if and only if the k functions $u^j x^j(t)$, where $u^1, \cdots, u^k$ denote k *arbitrary* real constants, are statistically independent in the additive sense.

In order to prove this theorem, put $x(t) = (x^1(t), \cdots, x^k(t))$ and $u = (u^1, \cdots, u^k)$. Since a distribution function and its Fourier-Stieltjes transform determine each other uniquely, one sees from (1) that (7) is equivalent to

$$
\begin{aligned}
L(u;\phi) &= \int_{R_x} \exp(iu\cdot x)\phi(dR_x) \\
&= \int_{R_{x^1}} \cdots \int_{R_{x^k}} \exp i\Big(\sum_{j=1}^{k} u^j x^j\Big)\phi^1(dR_{x^1}) \cdots \phi^k(dR_{x^k}) \qquad (9)\\
&= \prod_{j=1}^{k} \int_{R_{x^j}} \exp(iu^j x^j)\phi^j(dR_j) = \prod_{j=1}^{k} L(u^j;\phi^j).
\end{aligned}
$$

Since this is an identity in u, one can also say that (7) is equivalent to

$$L(\lambda u;\phi) = \prod_{j=1}^{k} L(\lambda u^j;\phi^j), \qquad (9\text{ bis})$$

where u is any vector and λ any scalar. Hence, on considering $u = (u^1, \cdots, u^k)$ as an arbitrary fixed vector and λ as the variable on a one-dimensional space R_λ, one sees from (4), from the M-criterion for the existence of an asymptotic distribution function (§ 2), and from the criterion (8) for a one-dimensional convolution $\psi = \psi^1 * \psi^2 * \cdots \psi^k$, that

(i) if it is assumed that the k scalar functions $x^j(t)$ are statistically independent, then the scalar function $u \cdot x(t)$ of t, where $u =$ Const., has an asymptotic distribution function, which is the convolution of the asymptotic distribution functions of the k scalar functions $u^j x^j(t)$ of t, where every $u^j =$ const.; and that

(ii) if it is assumed that the k scalar functions $u^j x^j(t)$ of t, where every $u^j =$ const., are statistically independent in the additive sense for arbitrary u^j, then the vector $x(t)$ and each of its components $x^j(t)$ have asymptotic distribution functions, ϕ and ϕ^j, which satisfy (9 bis), i. e. (9), hence also (7).

Since the assertion of (i) is precisely the assumption of (ii) and vice versa, the proof is complete.

8. On comparing (9) and (i)-(ii) with (4) and the M-criterion for the existence of an asymptotic distribution function (§ 2), one readily sees that the theorem proved in § 7 may be restated as follows:

k real, measurable, scalar functions $x^j(t)$ are statistically independent if and only if the limit $M\{\exp iu \cdot x(t)\}$ of $M_T\{\exp iu \cdot x(t)\}$ exists uniformly in every fixed sphere $|u| <$ const. and is such that

$$(10) \qquad M\{\exp iu \cdot x(t)\} = \prod_{j=1}^{k} M\{\exp iu^j x^j(t)\}$$

is an identity in u, where $x(t) = (x^1(t), \cdots, x^k(t))$ and $u = (u^1, \cdots, u^k)$.

This criterion for the statistical independence of k functions $x^1(t), \cdots, x^k(t)$ is due to Kac [3] in the particular case where the k functions $x^j(t)$ have the common period 1, i. e. for the case where the relative measure $\mu = \lim \mu_T$ on the infinite line $-\infty < t < +\infty$ reduces to the ordinary Lebesgue measure on the interval $0 \leqq t \leqq 1$. Since the latter measure is, while the relative measure is not, completely additive (and such as to preserve measurability when passing from two measurable sets to their common part), the approach used *loc. cit.*[3] in the case of periodic functions cannot be modified in such a way as to apply to the general case of relative measure. Methodically, the situation is that the proof given *loc. cit.*[3] in the particular case $M\{g(t)\} = \int_0^1 g(t)\,dt$ merely depends on the uniqueness theorem of Fourier-Stieltjes transforms; the Dirichlet discontinuity factors [3] merely serve the purpose of proving this uniqueness theorem.[16] On the other hand, the above treatment of the general case depends very much on the continuity theorem [9] of Fourier-Stieltjes transforms; a theorem essentially deeper than the uniqueness theorem (which is, of course, implied by it, and also by the inversion formula [10]).

9. For the case of relative measure, Kac and Steinhaus [4] have proposed a definition of the statistical independence of the k components $x^j(t)$ of a measurable vector function $x(t)$. This definition differs from ours for two reasons. In fact, *loc. cit.* [4]

(I) it is not required (as it is in § 4) that the limit (5) of (2) be a distribution function (cf. § 2 bis);

(II) it is required (as it is not in § 4) that the limit (5) of (2) should exist for every (and not "nearly all") intervals Q (cf. the example of Bohr,[7] mentioned in the Introduction).

On admitting (II), one is compelled [4] to impose smoothness conditions on the asymptotic distribution functions, even in case the functions $x^j(t)$ are bounded (in which case (I), of course, may be disregarded). Such smooth-

---

[16] Cf., e. g., E. K. Haviland, *American Journal of Mathematics*, vol. 56 (1934), pp. 625-658 (more particularly, pp. 638-641).

ness conditions, when satisfied, are often hard to establish and are not, in general, satisfied even for continuous periodic functions.

9 bis. No such additional smoothness condition is needed if, on disregarding (I) and (II), one defines statistical independence as in § 4. In fact, it follows from the M-criterion of the existence of an asymptotic distribution function (§ 2) and from (4), that k bounded measurable functions $x^j(t)$ are satistically independent if and only if any monomial of the $x^j(t)$ has an M-average and

$$(11) \qquad M\{\prod_{j=1}^{k}(x^j(t))^{n_j}\} = \prod_{j=1}^{k} M\{(x^j(t))^{n_j}\}$$

holds for k arbitrary non-negative integral exponents n_j. This becomes clear by observing [17] that, in case of bounded functions $x^j(t)$, the monomial averages (11), which are then identical with the respective momenta

$$(11\text{ bis}) \qquad \int_{R_x}\prod_{j=1}^{k}(x^j)^{n_j}\phi(dR_x) = \prod_{j=1}^{k}\int_{R_{x^j}}(x^j)^{n_j}\phi^j(dR_{x^j}),$$

simply are, up to the factors $i^{n_1+\cdots+n_k}n_1!^{-1}\cdots n_k!^{-1}$, the coefficients of $(u^1)^{n_1}\cdots(u^k)^{n_k}$ in the (convergent) expansion of (4) according to powers of the components u^j of u. Actually, the same holds also when the functions $x(t)$, $x^j(t)$, instead of being bounded, are such as to possess asymptotic distribution functions which belong to determined momentum problems.[18]

As to "very unbounded" cases, cases in which the momenta need not exist, cf. *loc. cit.* [5], p. 76.

10. Let, finally, be mentioned that the pathology of the relative measure $\mu = \lim \mu_T$ in the case of statistically independent functions appears to be essentially milder than in the general case.

First, if the k components $x^j(t)$ of $x(t)$ are statistically independent, then, by (5), (6) and (7), one has (for "nearly all" $Q \subset R_x$)

$$(7\text{ bis}) \qquad \mu[x(t) \subset Q] = \prod_{j=1}^{k}\mu[x^j(t) \subset Q^j], \quad (Q = Q^1 \times \cdots \times Q^k),$$

and vice versa. Since $[x(t) \subset Q]$ is the set of those points t which are common to the k sets $[x^j(t) \subset Q^j]$, it follows that the second of the two

---

[17] Cf. E. K. Haviland, *Proceedings of the National Academy of Sciences*, vol. 19 (1933), pp. 549-555.

[18] Cf. A. Wintner, *Mathematische Zeitschrift*, vol. 36 (1933), pp. 618-629; E. K. Haviland, *loc. cit.* [16]; H. Cramér and H. Woldt, *loc. cit.* [14].

pathological properties of the relative measure μ (cf. § 2) does not concern the present case.

The first of these properties, that concerning the failure of complete additivity, involves a more interesting and difficult question. In order to see this, let a sequence of measurable functions $\{x_n(t)\}$ be called convergent in relative measure on $-\infty < t < +\infty$ if, on placing $\bar{\mu} = \overline{\lim}_{T\to\infty} \mu_T$, one has

$$(12_1) \quad \bar{\mu}[\,|\,x_n(t) - x_m(t)\,| > \epsilon] \to 0,\ n, m \to \infty, \text{ for every fixed } \epsilon > 0;$$

and let $\{x_n(t)\}$ be said to tend to a limit function in relative measure if there exists a measurable function $x(t)$, $-\infty < t < +\infty$, such that

$$(12_2) \qquad \bar{\mu}[\,|\,x_n(t) - x(t)\,| > \epsilon] \to 0,\ n \to \infty, \text{ for every fixed } \epsilon > 0.$$

It is clear that if there exists such a limit function $x(t)$, it is uniquely determined by $\{x_n(t)\}$ save on a t-set of relative measure zero (in fact, $\bar{\mu}[\,|\,x(t) - y(t)\,| > \epsilon] = 0$ cannot hold for every $\epsilon > 0$, unless $x(t) = y(t)$ almost everywhere in relative measure). It is also clear that (12_2) implies (12_1); on the other hand, trivial examples show that (12_1) does not imply (as it does in the case of Lebesgue measure on a finite interval) the existence of an $x(t)$ satisfying (12_2).

Now, it seems to be a reasonable guess that if the sequence $\{x_n(t)\}$ is of the structure $x_n(t) = \xi_1(t) + \cdots + \xi_n(t)$, where $\{\xi_n(t)\}$ is an infinite sequence of functions such that the n functions $\xi_1(t), \cdots, \xi_n(t)$ are statistically independent for every n, then (12_1) implies the existence of an $x(t)$ satisfying (12_2). The truth of this conjecture follows in some cases from the theory of infinite convolutions [19]; while in another case, namely in the case of additive number-theoretical functions, the truth of the conjecture is suggested by a comparison of the sufficient criterion of Erdös [20] with the necessary and sufficient condition [21] for the convergence of an infinite convolution. But we succeeded in constructing an example which shows that the theorem is wrong for arbitrary $\{\xi_n(t)\}$.

QUEENS COLLEGE,
THE JOHNS HOPKINS UNIVERSITY.

---

[19] Cf. B. Jessen and A. Wintner, *loc. cit.* [5], Theorem 24 and § 16. The problem raised above is (roughly) to the effect, whether or not the converse of this Theorem 24 is true in case of finite sets which may be selected from the sequence $\{x_{n+1}(t) - x_n(t)\}$ consisting of statistically independent functions.

[20] P. Erdös, *Journal of the London Mathematical Society*, vol. 13 (1938), pp. 119-127.

[21] Cf. B. Jessen and A. Wintner, *loc. cit.* [5], Theorem 34.

THE QUATERNION CONGRUENCE $\bar{t}at \equiv b \pmod{g}$.*

By R. E. O'Connor and G. Pall.

1. Introduction. In this article we shall study the congruence

$$\bar{t}at \equiv b \pmod{g}, \tag{1}$$

where a and b are given pure quaternions, and t is a variable quaternion, but g is a rational integer. Special cases have been treated by Pall,[1] who considered mainly those properties needed for certain applications to the equation

$$h(8n+1) = x_1^2 + x_2^2 + x_3^2. \tag{1'}$$

We shall now determine for (1) necessary and sufficient conditions for solvability, in nearly all cases, and shall find the number and nature of the solutions. Interesting applications will be made, concerning the forms $x_1^2 + x_2^2 + x_3^2$ and

$$a \circ b = a_1b_1 + a_2b_2 + a_3b_3. \tag{2}$$

Notations. The letters $a, b, c, t, u, \cdots, z$ denote (Lipschitz) quaternions with rational integer coördinates (for an odd modulus g they may equally well be taken as Hurwitz quaternions); for example, $t = t_0 + i_1t_1 + i_2t_2 + i_3t_3$, the integer coördinates being distinguished by subscripts; a, b, c are pure, e. g. $a = i_1a_1 + i_2a_2 + i_3a_3$; $\bar{t} = t_0 - i_1t_1 - i_2t_2 - i_3t_3$; $Nt = \bar{t}t = t_0^2 + t_1^2 + t_2^2 + t_3^2$. We set $h = Na = a_1^2 + a_2^2 + a_3^2$. Except for the quaternion units i_α, letters with subscripts, as well as $f, g, h, \cdots, s$, denote rational integers; p an odd prime. It should be observed that for any t, $\bar{t}at$ is pure along with a.

As usual the question of solving (1) reduces to the case $g = p^n$ or 2^n (§ 14). On taking norms of both members of

$$\bar{t}at \equiv b \pmod{p^n}, \tag{3_n}$$

we obtain

$$(Nt)^2Na \equiv Nb \pmod{p^n}. \tag{4_n}$$

Hence the solvability of (3_n) requires the solvability for s of

* Received by the Editors, June 15, 1938.

[1] G. Pall, *American Journal of Mathematics*, vol. 59 (1937), pp. 895-913. Referred to as *AJ*.

$$(5_n) \qquad s^2Na \equiv Nb \pmod{p^n};$$

and the possible values of Nt are subject to

$$(6_n) \qquad (Nt)^2Na \equiv s^2Na \pmod{p^n}, \quad s \text{ a solution of } (5_n).$$

We shall find the striking result that:

(7) *if $p \nmid b$, there are equally many solutions t of* (3_n) *with Nt in any residue class* (mod p^n) *consistent with* (6_n).

But this does not hold if $p \mid b$: we shall treat this case only if $n = 1$, and then if $p \nmid a$, all solutions t will be found to satisfy $Nt \equiv 0 \pmod p$, even when $p \mid Na$. Our solution for $p \nmid b$, $n \geq 1$, will be complete.

The conditions obtained above on taking norms are sufficient for solvability (mod p^n) only if $p \nmid Na$. Let $p \mid Na$. If $p \nmid b$ the solvability of (3_n) will follow from that of (3_1) and the norm condition (5_n). In § 9, we prove

THEOREM 1. *Let a and b be pure, p an odd prime dividing Na and Nb but neither a nor b. Then* (3_1) *is solvable for t if and only if* $[a, b]_p = 1$, *where*

$$(8) \qquad \begin{aligned} [a, b]_p &= (2a \circ b \mid p) \quad \text{if } p \nmid a \circ b, \\ &= (m \mid p) \qquad \text{if } p \mid a \circ b, \end{aligned}$$

m being in the latter case an integer determined by $b \equiv ma \pmod p$.

The possibility of this is due to the interesting result, proved in § 8:

THEOREM 2. *Two vectors* (a_1, a_2, a_3), (b_1, b_2, b_3) *of norms divisible by p, are linearly dependent* (mod p) *if and only if* $p \mid a \circ b$.

Thus the symbol $[a, b]_p$ is defined, and equal to $+1$ or -1, for every pair of non-null vectors of norm zero (mod p). There are $p^2 - 1$ such vectors (mod p), and they form two mutually exclusive classes, which we will refer to as K and L, each containing $\frac{1}{2}(p^2 - 1)$ elements, and having the following property: according as a and b belong to the same or different classes K or L, (3_1) will be solvable or not, and $[a, b]_p$ will be $+1$ or -1. Properties of this symbol and of a similar one which tells whether the solutions of (1) with $g = 4$ or 8 satisfy $Nt \equiv 1$ or -1 (4), are developed in § 16. The following pretty property of Legendre symbols $(a \circ b \mid p)$ is proved therein:

THEOREM 3. *The value of $[a, b]_p$, defined in Theorem 1, is multiplied by* $(-1 \mid p)$ *if any coördinate a_1, a_2, or a_3 is changed in sign; and by* $(-2 \mid p)$ *if any two of a_1, a_2, a_3 are interchanged.*

By symmetry these operations may be applied to b. For example, let

$$a = 11i_1 + 4i_2 + 4i_3, \quad b = 10i_1 + 7i_2 + 2i_3, \quad p = 17.$$

Then $(a \circ b \mid p) = (146 \mid 17) = -1$, and $\bar{\imath}at \equiv b$ (17) is unsolvable. The value of $(a \circ b \mid p)$ remains -1, since $(-1 \mid p) = 1 = (-2 \mid p)$, when the coördinates of b are permuted or changed in sign; except that $17 \mid a \circ b$ when $\pm b = 2i_1 + 10i_2 - 7i_3$, and then $\pm b \equiv 11a$ (17), $(11 \mid 17) = -1$. But if

$$a = 7i_1 + 2i_2 + 2i_3, \quad b = 5i_1 - 4i_2 - 4i_3, \quad \text{and} \quad p = 19,$$

then $\bar{\imath}at \equiv b$ (19) is solvable with $t = 6$, since $b \equiv -2a$ (19) and $6^2 \equiv -2$; it is also solvable with b replaced by $-4i_1 + 5i_2 - 4i_3$, $4i_1 + 4i_2 + 5i_3$, etc., when $[a, b]_p = 1$; and not solvable with b replaced by $5i_1 + 4i_2 - 4i_3$, $4i_1 + 5i_2 - 4i_3$, etc., when $[a, b]_p = -1$.

Of three vectors a, b, c of norm zero (mod p), either two or three are in the same class K or L; hence follows

THEOREM 4. *Let p be an odd prime, and let no two of*

$$(9) \qquad a = (a_1, a_2, a_3), \quad b = (b_1, b_2, b_3), \quad c = (c_1, c_2, c_3)$$

be linearly dependent (mod p); *assume $p \mid Na$, Nb, and Nc. Then exactly one or three of the congruences*

$$(10) \qquad 2x_1^2 \equiv b_1c_1 + b_2c_2 + b_3c_3, \quad 2x_2^2 \equiv c_1a_1 + c_2a_2 + c_3a_3, \quad 2x_3^2 \equiv a_1b_1 + a_2b_2 + a_3b_3, \pmod{p}$$

are solvable in integers x_1, x_2, x_3.

The extension to $g = p^n$ $(n > 1)$ is accomplished in § 11 for $p \nmid b$. The modulus $g = 2^n$, which was treated extensively in AJ, is discussed further in §§ 12 and 13. Composite g and non-pure a and b are considered in §§ 14, 15.

A relation reminiscent of the connections between generic characters of quadratic forms is obtained in § 17 for the symbols $[a, b]_p$ and the corresponding symbol (mod 4 or 8). This enables us to give, in § 21, a much more satisfactory extension of Theorem 1 of AJ concerning the equation (1′), for any h such that

$$(11) \qquad h > 0, \quad h \neq 4n, \quad h \neq 8n + 7.$$

Sections 18-22 study the classes A and B of that theorem, and the distribution of values $[a, b]_p$.

2. THEOREM 5. *Let a and b be pure integral quaternions; p an odd prime not dividing $h = Na$. The congruence (3_1) is solvable for an integral quaternion t if and only if*

$$(Nb \mid p) = (Na \mid p) \text{ or } 0. \tag{12}$$

If $p \mid Nb$, (12) is evidently satisfied; then (3_1) has

$$\begin{array}{lll} & \text{the unique solution } t \equiv 0, & \text{if } b \equiv 0 \text{ and } (-Na \mid p) = -1, \\ (13) & 2p^2 - 1 \text{ solutions } t, & \text{if } b \equiv 0 \text{ and } (-Na \mid p) = 1, \\ & \theta = p - (-h \mid p) \text{ solutions } t, & \text{if } b \not\equiv 0. \text{ In all these cases } Nt \equiv 0. \end{array}$$

If $p \nmid Nb$ and (12) holds, we can choose an integer s satisfying (5_1); then (14) (3_1) has precisely 2θ solutions, θ of them satisfying each of $Nt \equiv \pm s$.

The necessity of (12) and of the respective restrictions $Nt \equiv 0$ or $\pm s$, follows from the discussion of (3_1) to (6_1).

Case I, $p \nmid Nb$. Then Nt is necessarily prime to p and we need only verify the following two statements:

(α) If, for a given a of norm prime to p, κ is the number of incongruent pure quaternion residues b (mod p) such that $(Nb \mid p) = (Na \mid p)$, and λ is the total number of residues t of norm prime to p, then $\lambda = 2\theta\kappa$.

(β) If (3_1) has a solution t, (14) holds.

Statement (α) is an immediate consequence of formulae for the number of solutions of quadratic congruences.[2] For given h and s, both prime to p, the number of solutions (b_1, b_2, b_3) of

$$b_1^2 + b_2^2 + b_3^2 \equiv s^2 h \pmod{p} \tag{15}$$

is exactly $\kappa' = p^2 + p(-h \mid p)$. The condition $(Nb \mid p) = (h \mid p)$ for determining b is equivalent to (15) with s^2 ranging over $\frac{1}{2}(p-1)$ quadratic residues. Hence $\kappa = \frac{1}{2}(p-1)\kappa'$. To evaluate λ we remark that, of the p^4 residues t (mod p), exactly $p^3 + p^2 - p$ satisfy $t_0^2 + t_1^2 + t_2^2 + t_3^2 \equiv 0$. Thus $\lambda = p^4 - p^3 - p^2 + p = p(p-1)^2(p+1)$. The quotient λ/κ is 2θ.

As regards statement (β) we can define a one-to-one correspondence between the solutions of (3_1) and those of

$$\bar{t}at \equiv a \pmod{p}. \tag{16}$$

---

[2] We employ specializations, where needed, of such formulae in Bachmann's *Zahlentheorie IV*, 1898, pp. 491-492. They are due to C. Jordan.

For if w is a particular solution of (3_1) and u is any other solution, $u\bar{w}/Nw$ $\pmod{p}$ is a solution of (16), having its norm congruent to $+1$ or $-1 \pmod{p}$ according as Nu is congruent to Nw or $-Nw$. Hence statement (β) follows from

LEMMA 1. *If a is pure and $p \nmid h$, the congruence* (16) *has precisely $\theta = p - (-h\,|\,p)$ solutions satisfying each of $Nt \equiv \pm 1 \pmod{p}$.*

Proof. If t satisfies (16) and $Nt \equiv 1$, then $at \equiv ta$ and $Nt \equiv 1$; and conversely. Since $p \nmid a$ the condition $at \equiv ta$, which expands into

$$(17) \qquad a_2t_3 \equiv a_3t_2, \quad a_3t_1 \equiv a_1t_3, \quad a_1t_2 \equiv a_2t_1 \pmod{p},$$

is equivalent to the existence of an integer λ satisfying

$$(18) \qquad t_1 \equiv \lambda a_1, \quad t_2 \equiv \lambda a_2, \quad t_3 \equiv \lambda a_3 \pmod{p}.$$

For any λ, t_1, t_2, t_3 are uniquely determined and t_0 is determined by

$$(19) \qquad t_0^2 + \lambda^2 h \equiv 1 \pmod{p},$$

to which the norm condition $Nt \equiv 1$ reduces. That (19) has θ solutions (t_0, λ) follows from Jordan's formulae (cf. preceding footnote).

The conditions (16) and $Nt \equiv -1$ reduce to the latter with $at \equiv -ta$, that is, $a_1t_1 + a_2t_2 + a_3t_3 \equiv 0$, $a_1t_0 \equiv a_2t_0 \equiv a_3t_0 \equiv 0 \pmod{p}$. Since $p \nmid a$ the last three are equivalent to $t_0 \equiv 0$. It remains to show that there are exactly θ sets $(t_1, t_2, t_3) \pmod{p}$ satisfying

$$(20) \qquad a_1t_1 + a_2t_2 + a_3t_3 \equiv 0, \quad t_1^2 + t_2^2 + t_3^2 \equiv -1, \pmod{p}.$$

We can suppose $p \nmid a_3$, solve (20_1) for t_3, $t_3 \equiv -e_1t_1 - e_2t_2$ where $e_1 \equiv a_1/a_3$, $e_2 \equiv a_2/a_3 \pmod{p}$; and substituting in (20_2) need to see that there are $p - (-h\,|\,p)$ solutions (t_1, t_2) of

$$(21) \qquad (1 + e_1^2)t_1^2 + 2(e_1e_2t_2)t_1 + \{(1 + e_2^2)t_2^2 + 1\} \equiv 0 \pmod{p}.$$

If $p \,|\, 1 + e_1^2$, (21) has no solutions with $t_2 \equiv 0$, and for each of the $p - 1$ residues t_2 prime to p, has an unique solution t_1. Further, $(-h\,|\,p) = 1$, since $p \,|\, a_3^2 + a_1^2$, $h \equiv a_2^2$, and $(-1\,|\,p) = 1$.

If $p \nmid 1 + e_1^2$, (21) has for each t_2, $1 + (\Delta\,|\,p)$ solutions t_1, where Δ is the discriminant of (21) considered as a quadratic in t_1. Now

$$\tfrac{1}{4}a_3^2\Delta \equiv -ht_2^2 + g, \text{ where } g \equiv a_2^2 - h \not\equiv 0 \pmod{p}.$$

Hence Δ is a quadratic residue for $\frac{1}{2}(p - 1 - (-h\,|\,p) - (gh\,|\,p))$ residues

t_2 and vanishes (mod p) for $1 + (gh|p)$ residues t_2; leading to θ values as required.

3. When $p|Nb$ we cannot reduce the problem by the artifice used above. Writing

$$(22)\quad at = -x_0 + i_1x_1 + i_2x_2 + i_3x_3 = \xi,\quad \bar{t}at = i_1c_1 + i_2c_2 + i_3c_3 = c,$$

as in AJ, § 11, we have

$$(23)\quad \begin{aligned} x_0 &= a_1t_1 + a_2t_2 + a_3t_3, & x_1 &= a_1t_0 + a_2t_3 - a_3t_2,\\ x_2 &= a_2t_0 + a_3t_1 - a_1t_3, & x_3 &= a_3t_0 + a_1t_2 - a_2t_1; \end{aligned}$$

$$(24)\quad c_1 = a_1(2(t_0^2 + t_1^2) - Nt) + 2a_2(t_0t_3 + t_1t_2) + 2a_3(-t_0t_2 + t_1t_3), \cdots, \cdots$$

the dots indicating that c_2 and c_3 are obtained from c_1 by cyclic permutation of subscripts 1, 2, 3, as x_2 and x_3 from x_1. Among the large number of identities true for any t and pure a the following will be useful:

$$(25_0)\quad x_0^2 + ht_0^2 = \tfrac{1}{2}(a_1c_1 + a_2c_2 + a_3c_3 + hNt),$$

$$(25_{123})\quad x_1^2 + ht_1^2 = \tfrac{1}{2}(a_1c_1 - a_2c_2 - a_3c_3 + hNt), \cdots, \cdots,$$

$$(26)\quad x_0x_1 - ht_0t_1 = \tfrac{1}{2}(a_2c_3 - a_3c_2), \cdots, \cdots,$$

$$(27)\quad x_2x_3 + ht_2t_3 = \tfrac{1}{2}(a_2c_3 + a_3c_2), \cdots, \cdots,$$

$$(28)\quad x_0t_1 + x_1t_0 = \tfrac{1}{2}(c_1 + a_1Nt), \cdots, \cdots,$$

$$(29)\quad x_2t_3 - x_3t_2 = \tfrac{1}{2}(c_1 - a_1Nt), \cdots, \cdots;$$

the dots indicating like formulae obtained cyclically; and $h = Na$.

Case II, $p|b$. By (3_1), $c \equiv b$ (mod p), and $Nt \equiv 0$; hence by (25), $ht_f^2 \equiv -x_f^2$ $(f = 0, 1, 2, 3)$. If $(-h|p) = -1$, the only solution is $t \equiv 0$. Let $(-h|p) = 1$. Since the right-members of (28)-(29) vanish, x_0, x_1, x_2, x_3 are proportional to $-t_0, t_1, t_2, t_3$ (mod p) respectively; that is

$$(30)\quad x_0 \equiv -\lambda t_0,\ x_1 \equiv \lambda t_1,\ x_2 \equiv \lambda t_2,\ x_3 \equiv \lambda t_3 \pmod p$$

for some integer λ. By (25), λ must be one of the two solutions of

$$(31)\quad \lambda^2 + h \equiv 0 \pmod p.$$

By (23) conditions (30) expand into

$$(30')\quad \lambda t_0 + a_1t_1 + a_2t_2 + a_3t_3 \equiv 0,\ a_1t_0 - \lambda t_1 - a_3t_2 + a_2t_3 \equiv 0, \cdots, \cdots.$$

All third order minor determinants in (30′) vanish in virtue of (31). Hence only two of the four congruences are independent. If p could divide every $\lambda^2 + a_a^2$ it would divide h. We can suppose $p \nmid \lambda^2 + a_1^2$; then the first two congruences (30′) are independent, t_2 and t_3 may be selected arbitrarily and then t_0 and t_1 are uniquely determined. Thus for each λ there are exactly p^2 solutions t, the only common solution for λ and $-\lambda$ being $t \equiv 0$. Now (13_2) will follow if the solutions t of (30′) with λ determined by (31) satisfy (3_1). To see this note that (30) holds for the x_f so obtained, $N(at) = \Sigma x_f^2 \equiv \lambda^2 Nt$, $hNt \equiv -hNt$ by (31), $Nt \equiv 0$, $\bar{t}at = \bar{t}\xi \equiv \bar{t}\lambda t = \lambda Nt \equiv 0 \pmod{p}$.

4. *Case* III, $p|Nb$, $p \nmid b$. We can suppose $p \nmid b_1$. Since $p|Nt$ by (3_1), and the left members of (28) or (29) differ from the coördinates of $\frac{1}{2}\bar{t}at$ by a multiple of Nt, the solutions of (3_1) coincide with those of either congruence triple

(32) $\quad x_0t_1 + x_1t_0 \equiv \frac{1}{2}b_1, \cdots, \cdots; \quad x_2t_3 - x_3t_2 \equiv \frac{1}{2}b_1, \cdots, \cdots;$

with $Nt \equiv 0$. Multiplying the second triple by t_1, t_2, t_3 and adding we get

$$b_1t_1 + b_2t_2 + b_3t_3 \equiv 0 \pmod{p}. \tag{33}$$

Eliminating x_0 between (32_2) and (32_3), and using (32_4), we get

$$b_1t_0 + b_3t_2 - b_2t_3 \equiv 0 \pmod{p}. \tag{34}$$

If t_0 and t_1 are determined from (33)-(34) then $Nt \equiv 0$. Replacing the x_i in either triple (32) by their values from (23) and substituting for t_0 and t_1 from (33)-(34), we find the three congruences obtained from

$$(a_3b_3 - a_1b_1 - a_2b_2)t_2^2 - 2(a_2b_3 + a_3b_2)t_2t_3 + (a_2b_2 - a_3b_3 - a_1b_1)t_3^2 \equiv \tfrac{1}{2}b_1^2 \tag{35}$$

when both sides are multiplied by b_1, b_2, b_3. It remains only to show that (35) has exactly θ solutions (t_2, t_3).

If $p \nmid a_3b_3 - a_1b_1 - a_2b_2$ the discriminant of (35) as a quadratic in t_2 is

$$\equiv 4b_1^2(-ht_3^2 + g) \pmod{p}, \text{ where } g = a_3b_3 - a_1b_1 - a_2b_2 \not\equiv 0.$$

The number of solutions is θ by the same argument as at the end of § 2.

If p divides both extreme coefficients in (35) then $p \nmid a_2b_3 + a_3b_2$, so that there are $p-1$ solutions. For then $p|a_1b_1$, $p|a_2b_2 - a_3b_3$, $p|a_1$, and since $p \nmid Na$, $p \nmid b_1$, and $p|Nb$, $p \nmid (a_2^2 + a_3^2)(b_2^2 + b_3^2)$,

$$p \nmid (a_2b_3 + a_3b_2)^2 = (a_2^2 + a_3^2)(b_2^2 + b_3^2) - (a_2b_2 - a_3b_3)^2.$$

But from $p|a_3b_3 - a_1b_1 - a_2b_2$ and $p|Nb$ follows $(-h|p) = 1$. For we have

LEMMA 2. *Let* $h = a_1^2 + a_2^2 + a_3^2$, $(-h|p) = -1$. *Then the congruences*

$$(36) \qquad a_1k_1 + a_2k_2 + a_3k_3 \equiv 0, \quad k_1^2 + k_2^2 + k_3^2 \equiv 0, \ (\text{mod } p)$$

have no solution k_1, k_2, k_3 *other than* $k_1 \equiv k_2 \equiv k_3 \equiv 0$.

The proof reduces to a discussion of (21) with the last $+1$ replaced by 0; the discriminant of the quadratic is $-4hk_2^2/a_3^2$, whence the lemma follows and the proof of Theorem 5 is complete.

5. The norm-class (mod *n*) of *t*. By this we mean the residue class (mod n) to which Nt belongs. Theorem 5 can be paraphrased succinctly thus:

COROLLARY 1. *For pure* a, b, Na *being prime to* p, (3_1) *has precisely* η *solutions in each norm-class* (mod p) *consistent with* (4_1); η *is* 1, $2p^2-1$, *or* $p-(-Na|p)$ *according as*

$$p|b \text{ and } (-Na|p) = -1, \quad p|b \text{ and } (-Na|p) = -1, \quad \text{or } p \nmid b.$$

The remainder of Theorem 5 may be deduced by solving (5_1) for s.

6. THEOREM 6. *Let* a *and* b *be pure, integral quaternions,* p *an odd prime,* $p|Na$, $p \nmid a$. *Assume that* (3_1) *has at least one integral solution* $t = u$. *Then if* $p|b$, (3_1) *has exactly* p^2 *solutions* t, *all satisfying* $Nt \equiv 0$. *If* $p \nmid b$, (3_1) *has precisely* $2p^2$ *solutions* t, *of which* $2p$ *satisfy each of*

$$(37) \qquad Nt \equiv r \ (\text{mod } p), \quad r = 0, 1, \cdots, p-1.$$

Except for the distribution of norm residues given by (37) this was proved in *AJ* (Theorem 11), by demonstrating the equivalence, if $p|Na$, of

$$(38) \qquad \bar{t}at \equiv \bar{u}au \ (\text{mod } p) \ \text{ and } \ at \equiv \pm au \ (\text{mod } p).$$

This equivalence can be deduced more simply as follows. Set

$$au = -y_0 + i_1y_1 + i_2y_2 + i_3y_3, \quad \bar{u}au = c',$$

as in (22). Then (38_1) becomes $c \equiv c'$ and implies by (25)-(27) that $x_fx_g \equiv y_fy_g$ $(f, g = 0, 1, 2, 3)$, whence $x_f \equiv \pm y_f$ $(f = 0, 1, 2, 3)$ with all the signs alike. Conversely the latter conditions, which are the same as (38_2), imply by (25)-(27) that $a_fc_g \equiv a_fc'_g$ $(f, g = 1, 2, 3)$, $c \equiv c'$.

As in the discussion of (66) in *AJ*, $at \equiv \epsilon au$ (where $\epsilon = \pm 1$) when expanded reduces to two independent congruences, for example if $p \nmid a_1$ to

$$(39)\qquad \begin{aligned} a_1t_1 + a_2t_2 + a_3t_3 &\equiv k_1, & k_1 &= \epsilon(a_1u_1 + a_2u_2 + a_3u_3),\\ a_1t_0 - a_3t_2 + a_2t_3 &\equiv k_2, & k_2 &= \epsilon(a_1u_0 - a_3u_2 + a_2u_3),\end{aligned}$$

the other congruences $x_2 \equiv \epsilon y_2$, $x_3 \equiv \epsilon y_3$, being linear combinations of these. Clearly k_1 and k_2 will be both zero $(\bmod p)$ if and only if $au \equiv 0$, and hence by a special case of (38) if and only if $\bar{u}au \equiv 0$. By (39),

$$(40)\qquad \begin{aligned} t_0^2 + t_1^2 + t_2^2 + t_3^2 \equiv{} & (k_1^2 + k_2^2)/a_1^2 \\ & - 2t_2(k_1a_2 - k_2a_3)/a_1^2 - 2t_3(k_1a_3 + k_2a_2)/a_1^2. \end{aligned}$$

In case $p\,|\,b$, $k_1 \equiv k_2 \equiv 0$, t_2 and t_3 may be chosen arbitrarily and t_0, t_1 are then fixed by (39); hence there are p^2 solutions t and all have $Nt \equiv 0$ in view of (40). If $p \nmid b$, consider (39) with a fixed sign, and (40). Let r be a given integer. Then $Nt \equiv r \pmod{p}$ has precisely p solutions t_2, t_3 and thus (39) has precisely p solutions t with $Nt \equiv r$, unless simultaneously

$$k_1a_2 - k_2a_3 \equiv 0, \quad k_1a_3 + k_2a_2 \equiv 0, \pmod{p}.$$

But this would imply, since k_1 and k_2 are not both zero, that $a_2^2 + a_3^2 \equiv 0$, whence $p\,|\,a_1$. Since k_1 or k_2 is altered with the sign of ϵ, the p solutions obtained with $\epsilon = 1$ are distinct from those with $\epsilon = -1$.

As shown on p. 911 of *AJ* it follows that exactly half the $p^2 - 1$ residues $b \pmod{p}$ such that $Nb \equiv 0$ but $b \not\equiv 0$ are represented by $\bar{t}at$, for a given a such that $p\,|\,Na$ but $p \nmid a$, each being represented for $2p^2$ residues t. The following corollary will not be used in this article but is useful in finding rapidly the set of residues $\bar{t}at \pmod{p}$.

COROLLARY 2. *If $p\,|\,Na$ and $p \nmid a_1$, the $\frac{1}{2}(p^2 - 1)$ non-zero residues $(\bmod p)$ represented by $\bar{t}at$ are obtained each twice when t is given the $p^2 - 1$ values $t_0 + i_1t_1$, $t_0 = 0, 1, \cdots, p-1$, $t_1 = 0, 1, \cdots, p-1$, omitting $t_0 = t_1 = 0$.*

For by the discussion of (39) each non-zero residue b is represented by $\bar{t}at$ for two values t with $t_2 = t_3 = 0$. Evidently if $p \nmid a_2$ or a_3, the same result holds with $t_0 + i_2t_2$, $t_0 + i_3t_3$ in place of $t_0 + i_1t_1$.

7. LEMMA 3. *Let p be an odd prime, r an integer, a_1, a_2, a_3 integers not all divisible by p, $p\,|\,a_1^2 + a_2^2 + a_3^2$. Then the congruences*

$$(41)\qquad x_1^2 + x_2^2 + x_3^2 \equiv 0, \quad a_1x_1 + a_2x_2 + a_3x_3 \equiv r, \pmod{p}$$

have precisely p incongruent, integral sets of solutions (x_1, x_2, x_3).

We can adjust subscripts so that $p \nmid a_2a_3$. Then (41_2) can be solved for x_3, and the result substituted in (41_1), yielding

$$(42) \qquad (1+e_1^2)x_1^2 - 2e_1(s-e_2x_2)x_1 \\ + (1+e_2^2)x_2^2 - 2e_2sx_2 + s^2 \equiv 0 \pmod p,$$

where $s \equiv r/a_3$, $e_1 \equiv a_1/a_3$, $e_2 \equiv a_2/a_3$, $x_3 \equiv s - e_1x_1 - e_2x_2$. Since $p \nmid 1 + e_1^2$, (42) has for each value x_2, $1 + (\Delta|p)$ solutions x_1, where Δ is the discriminant of (42) taken as a quadratic in x_1. Using $p | 1 + e_1^2 + e_2^2$ we find

$$\Delta \equiv 4(2x_2e_2s - s^2) \pmod p, \text{ where } p \nmid e_2.$$

If $p|s$, Δ vanishes (mod p) for each of the p incongruent choices of x_2; if $p \nmid s$, Δ is a quadratic residue for $\frac{1}{2}(p-1)$ choices x_2, and vanishes for one; in either case there are p solutions.

8. Proof of Theorem 2. By lemma 3 with $r = 0$, when a is given there are exactly p residues b such that $p | Nb$ and $p | a \circ b$. Since $a, 2a, \cdots, pa$ (mod p) exhaust these possibilities, no two of them being congruent, Theorem 2 follows.

9. Proof of Theorem 1. For a given a, precisely $\frac{1}{2}(p^2-1)$ residues b, non-zero mod p satisfy $[a, b]_p = 1$. For, by lemma 3, there are p solutions b of

$$a_1b_1 + a_2b_2 + a_3b_3 \equiv r \pmod p$$

for each of the $\frac{1}{2}(p-1)$ values r such that $(2r|p) = 1$, and besides we have $b \equiv ma$ with m one of $\frac{1}{2}(p-1)$ quadratic residues of p; in all

$$\tfrac{1}{2}p(p-1) + \tfrac{1}{2}(p-1) = \tfrac{1}{2}(p^2-1) \text{ residues } b.$$

That these $\frac{1}{2}(p^2-1)$ residues b coincide with the $\frac{1}{2}(p^2-1)$ non-zero residues represented by $\bar{t}at$ will be evident when we show that $[a, b]_p = 1$ is a necessary condition for the solvability of (3_1). This follows from (25), where now $h \equiv 0$ and $c \equiv b$: clearly $2a \circ b$ must be a quadratic residue of p unless $a \circ b \equiv 0$; and then by Theorem 2, $b \equiv ma$ for some integer m; hence replacing the c_f in (25) by ma_f, we have $ma_f^2 \equiv x_f^2$ for $f = 1, 2, 3$, that is $(m|p) = 1$.

COROLLARY 3. *If p does not divide a, b, but divides their norms, then* (3_1) *is solvable only if $[a, b]_p = 1$, and then has precisely $2p$ solutions in each norm-class consistent with the norm condition* (4_1). (Cf. § 5.)

10. THEOREM 6. *Let a and b be pure integral quaternions, p an odd*

prime, $p \nmid b$. Let η denote the number of solutions t of (3_1) *in each norm-class* $\pmod{p}$ *consistent with* (4_1). *Then* (3_n) *has precisely ηp^{n-1} solutions* $t \pmod{p^n}$ *in every norm-class* $\pmod{p^n}$ *consistent with* (4_n).

Remarks. The values of η are given in corollaries 1 and 3. By solving the quadratic congruence (5_n), by familiar methods, we can derive from Theorem 6 complete information, when $p \nmid b$, regarding the number of solutions of (3_n) and their distribution in norm-classes. For example:

Let $(Na|p) = (Nb|p) \neq 0$. *Then* (3_n) *is solvable for every integer* $n \geqq 1$, *and has precisely $2\theta p^{n-1}$ solutions* $t \pmod{p^n}$, *half of them satisfying each of* $Nt \equiv s$ *or* $-s \pmod{p^n}$, *s satisfying* (5_n). *Here by Theorem 5,* $\theta = p - (-Na|p)$.

Let p divide Nb but not a nor b. Let p^{r_b} be the least power of p not dividing Nb. Assume that (3_1) *is solvable. Then if* (5_n) *has a solution s when* $n = r_b$, (3_n) *is solvable for every n. But if* (5_n) *is not solvable when* $n = r_b$, (3_n) *is solvable when* $n < r_b$, *and not when* $n \geqq r_b$. This follows from the fact that for $n < r_b$, (5) is trivially solvable; but is solvable with $n \geqq r_b$ if and only if it is solvable when $n = r_b$.

11. Proof of Theorem 6. Let $n > 1$ and denote an arbitrary solution of (3_{n-1}) by t, and of (3_n) by w. Since every w is a t, we obtain each w $\pmod{p^n}$ exactly once in $w = t + p^{n-1}u$, at t ranges over all solutions of $(3_{n-1}) \pmod{p^{n-1}}$, and for each t, u ranges over all solutions of

$$(\bar{t} + p^{n-1}\bar{u})a(t + p^{n-1}u) \equiv b \pmod{p^n}. \tag{43}$$

For u is hereby determined to modulus p, since (43) reduces to

$$\bar{t}au + \bar{u}at \equiv -v \pmod{p}, \tag{44}$$

where $\bar{t}at - b = p^{n-1}v$. Set $v = i_1v_1 + i_2v_2 + i_3v_3$ and use (22); (44) becomes

$$\begin{aligned} x_1u_0 + x_0u_1 - x_3u_2 + x_2u_3 &\equiv -\tfrac{1}{2}v_1, \\ x_2u_0 + x_3u_1 + x_0u_2 - x_1u_3 &\equiv -\tfrac{1}{2}v_2, \\ x_3u_0 - x_2u_1 + x_1u_2 + x_0u_3 &\equiv -\tfrac{1}{2}v_3, \pmod{p}. \end{aligned} \tag{45}$$

The matrix of coefficients of u_0, u_1, u_2, u_3 has determinants equal to

$$\pm x_0N\xi, \ \pm x_1N\xi, \ \pm x_2N\xi, \ \pm x_3N\xi \text{ respectively.} \tag{46}$$

The case $p \nmid NaNb$ is easy. Then $p \nmid Nt$, whence $p \nmid N\xi = NaNt$, and the numbers (46) are not all zero $\pmod{p}$. Hence one of the u_i may be

chosen arbitrary (mod p), the others then uniquely determined. Hence (3_n) has p times as many solutions as (3_{n-1}), there being p values u (mod p) for each t. Lemma 4 follows for $p \nmid Nb$.

Let $p \mid Nb, p \nmid b$. Then $p \mid N\xi$, and the matrix of coefficients in (45) is of rank < 3. As shown in AJ in discussing (58), the linear combination of (45) with multipliers c_1, c_2, c_3 vanishes. The condition

$$(47) \qquad c_1v_1 + c_2v_2 + c_3v_3 \equiv 0 \pmod p$$

is thus necessary to the solvability of (45). Since $c - b = p^{n-1}v$,

$$Nb = Nc - p^{n-1}(\bar{c}v + \bar{v}c) + p^{2n-2}\bar{v}v, \quad \bar{c}v + \bar{v}c = 2(c_1v_1 + c_2v_2 + c_3v_3).$$

Hence (47) is equivalent to $p^n \mid Nb - Nc$, which is (4_n). A solution t of (3_{n-1}) will therefore not provide solutions w of (3_n) unless (4_n) holds; if (4_n) holds we call t *eligible*. If t is eligible and $p \nmid c_f$ for $f = 1, 2,$ or 3, then the f-th congruence (45) is a linear combination of the other two.

To compare the norms of w and t we have

$$Nw = N(t + p^{n-1}u) = Nt + p^{n-1}(\bar{t}u + \bar{u}t) + p^{2n-2}\bar{u}u,$$
$$Nw \equiv Nt + 2p^{n-1}(t_0u_0 + t_1u_1 + t_2u_2 + t_3u_3), \pmod{p^n}.$$

Thus $Nw \equiv Nt \pmod{p^{n-1}}$, and Nw can be made to have any residue (mod p^n) consistent with this by choice of μ in

$$(48) \qquad t_0u_0 + t_1u_1 + t_2u_2 + t_3u_3 \equiv \mu \pmod p.$$

Consider now the system of three congruences (48) and (45_2) and (45_3), under the hypotheses that $p \mid Nb$, t is eligible, and $p \nmid c_1$ (which we can suppose if $p \nmid b$), whence (45) does reduce to its last two congruences. The matrix of coefficients of u_0, u_1, u_2, u_3 has determinants equal respectively to

$$c_1x_0, \ c_1x_1, \ c_1x_2, \ c_1x_3, \text{ apart from sign.}$$

This can be seen conveniently as follows. Denote by Δ_i $(i = 0, 1, 2, 3)$ the determinant of order 3 formed by dropping the coefficients of u_i; since

$$c = \bar{t}\xi = (t_0 - i_1t_1 - i_2t_2 - i_3t_3)(-x_0 + i_1x_1 + i_2x_2 + i_3x_3)$$

is a pure quaternion, we have

$$0 = -t_0x_0 + t_1x_1 + t_2x_2 + t_3x_3 = c_0, \quad c_1 = t_0x_1 + t_1x_0 - t_2x_3 + t_3x_2;$$

but, identically in components of t and ξ,

$$\Delta_0 - c_1x_0 = c_0x_1, \qquad \Delta_1 - c_1x_1 = - c_0x_0,$$
$$\Delta_2 + c_1x_2 = - c_0x_3, \qquad \Delta_3 - c_1x_3 = - c_0x_2.$$

Hence the statement concerning the matrix is accurate. Since p cannot divide ξ without dividing c, the matrix is of rank 3: for any assigned integer μ there are precisely p solutions of the system (45) and (48).

Noting, finally, that (3) is unsolvable if $p \mid Na$ but $p \nmid Nb$, we have

LEMMA 4. *If t is a solution of (3_{n-1}) and $n > 1$, then the congruence*

$$\bar{w}aw \equiv b \pmod{p^n} \tag{49}$$

has precisely p solutions of the form $w = t + p^{n-1}u$, if $p \nmid Nb$; if $p \mid Nb$, (49) has solutions of this form if and only if t is an eligible solution of (3_{n-1}), that is t satisfies also (4_n), and then (49) has precisely p^2 solutions w of this form, p of them satisfying each of

$$Nw \equiv Nt + p^{n-1}\nu \pmod{p^n}, \qquad (\nu = 0, 1, \cdots, p-1). \tag{50}$$

In case $p \nmid Nb$, represent the two solutions of (5_n) by s and $-s$. These may be taken to represent the two solutions of (5_{n-1}). If t belongs to the norm-class $s \pmod{p^{n-1}}$, the p derived solutions $w = t + p^{n-1}u$ evidently belong only to the norm-class $s \pmod{p^n}$. For $Nw \equiv Nt$, $s \not\equiv -s \pmod{p}$. This proves the case $p \nmid Nb$ of

LEMMA 5. *If (3_{n-1}) has precisely k solutions $\pmod{p^{n-1}}$ in each norm-class satisfying (4_{n-1}), then (3_n) has precisely kp solutions $\pmod{p^n}$ in each norm-class satisfying (4_n).*

Let $p \mid Nb$. Then $p \mid NaNt$. Hence if t and t' are solutions of (3_{n-1}) such that $Nt \equiv Nt' \pmod{p^{n-1}}$, $Na(Nt')^2 \equiv Na(Nt)^2 \pmod{p^n}$. That is all or none of the solutions t in a norm-class $\pmod{p^{n-1}}$ are eligible. We may speak of an *eligible norm-class*. Every t in an eligible norm-class $\pmod{p^{n-1}}$ leads to the same set of p norm-classes (p^n), given by (50); and evidently no t'' in a different norm-class can lead to any of these. This proves that every norm-class $\pmod{p^n}$ which is derived from an eligible norm-class $\pmod{p^{n-1}}$ contains precisely kp solutions w. The lemma will follow when we show that every norm-class $\pmod{p^n}$ satisfying (4_n) is derivable from an eligible norm-class $\pmod{p^{n-1}}$. Let s satisfy (5_n) and hence (5_{n-1}). By the hypothesis of the lemma there is a t of norm $s \pmod{p^{n-1}}$ (in fact k of them). Since $p \mid NaNt$, $(Nt)^2Na \equiv s^2Na \equiv Nb \pmod{p^n}$; so that t is an eligible solution of (3_{n-1}). Clearly s will be among the norm-classes $\pmod{p^n}$, as in (50), to which t leads.

12. Modulus 2^n. Let a and b denote pure quaternions, neither divisible by 2. Set $l = 1$ or 2 according as Na is odd or $\equiv 2 \pmod 4$. It is shown in AJ that if $n \geqq 2$, the congruence

$$\bar{t}at \equiv b \pmod{l2^n} \tag{51}$$

is solvable for t if and only if the congruence

$$\bar{t}at \equiv b \pmod{4l} \tag{52}$$

is solvable; that (52) is solvable if and only if

$$a \equiv b \pmod 2, \quad Na \equiv Nb \pmod{8l}. \tag{53}$$

If $n = 1$ and $l = 2$, (51) is solvable if and only if

$$a \equiv b \pmod 2, \quad Na \equiv Nb \pmod 8; \tag{54}$$

and if $l2^n = 2$, if and only if $a \equiv b \pmod 2$.

The condition, similar to (4), obtained on taking norms (AJ § 3) is

$$Na(Nt)^2 \equiv Nb \pmod{l2^{n+1}}. \tag{55}$$

This implies that in all solutions of (51),[3]

$$Nt \equiv s \text{ or } Nt \equiv -s \pmod{2^n}, \tag{56}$$

where s denotes a solution of

$$s^2 Na \equiv Nb \pmod{l2^{n+1}}. \tag{57}$$

But, *unlike the case of an odd modulus, all solutions t satisfy only one of conditions* (56). (AJ, Th. 3.)

Assuming (53), we define $\epsilon(a, b) = +1$ or -1 according as

$$Nt \equiv 1 \text{ or } Nt \equiv -1 \pmod 4 \tag{58}$$

for the solutions of (52). The value of $\epsilon(a, b)$ is given in a practicable form in AJ, Theorem 5. See also § 16.

---

[3] A correction should be made in AJ, Theorem 6, p. 901. Formula (18) should be replaced by our present (56)-(57); and on p. 902, line 13, $\bar{t}t \equiv \pm 1$ should be changed to $\bar{t}t \equiv \pm s$. Further corrections: In (42), change the second $+$ to $-$. On p. 908, line 8 from bottom, $\bar{t}au$ should be $\bar{t}\bar{a}u$. On p. 913, line 6, change Nv to Nt. A desirable amplification of AJ, § 15 is included in our present § 21.

12a. We omit the verification that the following more complicated, but possibly more significant, expressions for $\epsilon(a, b)$ agree with AJ Theorem 5; but note that the verification for Na even is much simplified by use of the identity

$$M \pm \tfrac{1}{2}K = \tfrac{1}{4}\{(a_1 \pm b_1)^2 + (a_2 \pm b_2)^2\} + \tfrac{1}{4}(a_3 \pm b_3)^2,$$

where

$$M = \tfrac{1}{4}(Na + Nb), \quad K = a \circ b = a_1b_1 + a_2b_2 + a_3b_3.$$

If $K \equiv 2 \pmod 4$, denote by k the residue (mod 16) of $\frac{1}{2}K$ between -8 and $+8$, by $(-)_k$ the sign of k. Then Theorem 5 of AJ implies the following:

$$\text{if } Na \equiv 3 \pmod 8,\ \epsilon(a, b) = (-1)^{(K+1)/2};$$
$$\text{if } Na \equiv 1 \pmod 4,\ \epsilon = (-1)^{(Na-Nb)/8}(2|K)(2|Na);$$
$$\text{if } Na \equiv 2(8),\ \epsilon = (-1)^{(M-1)/4+K/8} \text{ if } 4|K, \text{ otherwise } \epsilon = (-1)^{(|k|-1)/2}\delta,$$

where $\delta = (-1)^{(M-1)/8}$ if $M \equiv 1$, $\delta = (-1)^{(M+3)/8}(2|k)$ if $M \equiv 5(8)$;

$$\text{if } Na \equiv 6(8),\ \epsilon = (-1)^{(M-3)/4+(K+4)/8} \text{ if } 4|K, \text{ otherwise } \epsilon = (-)_k\delta,$$

where $\delta = (-1)^{(M-3)/8}$ if $M \equiv 3$, $\delta = (-1)^{(M-7)/8}(2|k)$ if $M \equiv 7(8)$. Note that if $Na \equiv 2(4)$, K is even but cannot be $\equiv Na + 2(8)$.

13. The determination of the number of solutions of (51) is comparatively easy, for a and b restricted as above. Indeed, Nt being necessarily odd,

(59) (51) *has the same number of solutions t for every b for which it is solvable.*

For if u is one solution, the number of solutions is the same as that of $\bar{t}at \equiv \bar{u}au$, hence of $\bar{w}aw \equiv a \pmod{l2^n}$, where $w = (t\bar{u})/Nu$ is, with t, an independent integral quaternion variable (mod $l2^n$).

Hence by merely counting the number of residues b (mod $l2^n$) satisfying (53) (or if $n = 1$, $l = 2$, (54)), and dividing into the total number of residues t of odd norm, we obtain the number of solutions t for any solvable b.

We may restrict t to modulus $l2^{n-1}$ (if $n > 1$). For, if t is determined to modulus $2^r (r > 0)$, $\bar{t}at$ is determined to modulus 2^{r+1}:

$$(\bar{t} + 2^r\bar{v})a(t + 2^rv) = \bar{t}at + 2^r(\bar{t}av - \overline{v}\bar{a}t) + 2^{2r}\bar{v}v,$$

where the quaternion $\bar{t}av$ minus its conjugate is an even integer.

Proceeding in this way we find the number of solutions of (51) is

(59′) 2^{n-2} if $Na \equiv 3\ (8)$, 2^{n-1} if $Na \equiv 1\ (4)$, 2^{n+1} if $Na \equiv 2\ (4)$;

t being taken (mod $l2^{n-1}$) and $n \geqq 2$, (53) assumed. If $l2^n = 2$ and $a \equiv b(2)$, the number of t's (mod 2) is 8; if (54) holds, $n = 1$, and $l = 2$, the number of t's (mod 2) is 2.

14. Extension to general odd modulus. If m_1 and m_2 are odd coprime integers and t', t'' are solutions of $\bar{t}at \equiv b$ (m_1) and (m_2) respectively, there is an unique solution t (mod m_1m_2) of the congruences $t \equiv t'$ (m_1) and $t \equiv t''$ (m_2), and this is a solution of $\bar{t}at \equiv b$ (m_1m_2). It is easily seen that there is a one-to-one correspondence between the pairs (t', t'') and the quaternions t; also between the pairs of norm-classes $\mathfrak{n}_1$ (say) (mod m_1), $\mathfrak{n}_2$ (mod m_2), and the norm-classes $\mathfrak{n}$ (mod m_1m_2). If there are n_1 solutions t' (m_1) in the norm-class $\mathfrak{n}_1$ and n_2 solutions t'' (m_2) in the norm-class $\mathfrak{n}_2$, there are precisely n_1n_2 solutions t (m_1m_2) in the norm-class $\mathfrak{n}$ corresponding to $(\mathfrak{n}_1, \mathfrak{n}_2)$. An easy application of these ideas gives:

THEOREM 7. *Let $g = p_1^{r_1} p_2^{r_2} \cdots p_s^{r_s}$ where the p_i are all different odd primes. Let a and b be pure integral quaternions divisible by none of the p_i. Let η_i $(i = 1, 2, \cdots, s)$ be the number of solutions of $\bar{t}at \equiv b$* (mod p_i) *in each norm-class consistent with* $(Nt)^2Na \equiv Nb$ (mod p_i). *Then the congruence* (1) *has precisely*

$$\eta_1\eta_2 \cdots \eta_s p_1^{r_1-1} \cdots p_s^{r_s-1}$$

solutions t (mod g) *in each norm-class* (mod g) *consistent with*

$$(Nt)^2Na \equiv Nb \ (g).$$

The η_i can be evaluated by the criteria given above and the number of solutions of (1) readily determined.

15. Extension to non-pure quaternions. Preserving the notation heretofore used, the problem of solving

$$\bar{t}(a_0 + a)t \equiv b_0 + b \pmod{g} \tag{60}$$

is equivalent to that of solving

$$\bar{t}at \equiv b, \quad a_0Nt \equiv b_0 \pmod{g}; \tag{60'}$$

and hence to finding the solutions of $(60'_1)$ in norm-classes consistent with $(60'_2)$. In particular, *if g is odd and* $(60'_1)$ *is solvable, then* (60) *is solvable if and only if*

$$r^2Na \equiv Nb, \quad ra_0 \equiv b_0 \pmod{g}$$

have a common solution r. A necessary condition is that

$$a_0^2 Nb \equiv b_0^2 Na \pmod{g}.$$

Using Theorem 5 and considering the cases $p | Nb$ and $p \nmid Nb$, we have

THEOREM 8. *If g is an odd prime p; $p \nmid Na$, $p \nmid a_0$; then* (60) *is solvable for t if and only if*

$$a_0^2 Nb \equiv b_0^2 Na \pmod{p}.$$

16. Properties of $[a, b]_p$ and $\epsilon(a, b)$. The symbol $[a, b]_p$ defined in Theorem 1, evidently has the following properties, for any non-zero, pure quaternions a, b, c of norm zero (mod p):

$$(61) \qquad [a, a]_p = 1, \quad [a, b]_p = [b, a]_p, \quad [a, b]_p = [a, c]_p [b, c]_p,$$
$$[a, \lambda b]_p = (\lambda | p) \cdot [a, b]_p \text{ for any integer } \lambda \text{ prime to } p.$$

When $t = i_1$, $\bar{t}at = i_1 a_1 - i_2 a_2 - i_3 a_3$; thus the signs of two coördinates of a may be changed without changing its class K or L (cf. sequel to Th. 2). Hence if a' is obtained from a by changing the sign of one coördinate,

$$[a', b]_p = [-a, b]_p = (-1 | p)[a, b]_p.$$

When $t = i_1 + i_2$, $\bar{t}at = 2a^\dagger$, where $a^\dagger = i_1 a_2 + i_2 a_1 - i_3 a_3$; set $a'' = i_1 a_2 + i_2 a_1 + i_3 a_3$; then

$$[a, b'']_p = (2 | p)[a, 2b'']_p = (-2 | p)[a, 2b^\dagger]_p = (-2 | p)[a, b]_p.$$

These two properties are summarized in Theorem 3.

The symbol $\epsilon(a, b)$ is defined only for pure a, b, neither divisible by 2, and satisfying (53). From Lemma 1 of *AJ* (p. 904-5) we immediately draw the following conclusions:

$$(62) \qquad \epsilon(a, a) = 1, \quad \epsilon(a, b) = \epsilon(b, a), \quad \epsilon(a, b) = \epsilon(a, c) \cdot \epsilon(b, c),$$

$$(63) \qquad \epsilon(a', b) = (-1 | H) \cdot \epsilon(a, b), \quad \epsilon(a'', b) = (-2 | H) \cdot \epsilon(a, b);$$

$$(64) \qquad \text{if } l = 2, \quad \epsilon(a, 5b) = -\epsilon(a, b).$$

Here a' is obtained from a by changing the sign of one coördinate, a'' is obtained from a by interchanging two coördinates of like parity, $Na = lH$, $l = 1$ or 2. Since a and b are only determined for this symbol to modulus $4l$, these formulae provide complete data on the effects of interchanging coördinates, changing their signs, or multiplying by an odd integer. We easily verify that if m is prime to $2h$,

$$(64') \qquad \epsilon(a, mb) = (-1)^{\frac{1}{2}(m-1) \cdot \frac{1}{2}(H-1)} (l | m) \cdot \epsilon(a, b) = (h | m)(m | H) \epsilon(a, b).$$

17. Two proper representations of an integer h as a sum of three squares have the following property:

THEOREM 9. *Let a and b be proper, pure quaternions of norm h, $a \equiv b(2)$; set $h = lp_1 \cdot \cdot \cdot p_s$, where $l = 1$ or 2 and the p_i are odd primes. Then*

$$\prod_{i=1}^{s} [a, b]_{p_i} = \epsilon(a, b). \tag{65}$$

In the proof we use the following lemma from an article by Pall, "On the Arithmetic of Quaternions," Lemma 17:

LEMMA 6. *For a and b satisfying the hypotheses of Theorem 9, there exists a proper quaternion t of norm m prime to $2h$, such that*

$$\bar{t}at = mb', \tag{66}$$

where b' is obtained from b by permutations and sign-changes of b_1, b_2, b_3.

Since $\bar{t}at \equiv (Nt)a \pmod 2$ (AJ (4)), (66) implies that $a \equiv b'(2)$. From (63) and Theorem 3 follows that if (65) holds with b' in place of b it is true as stated. For a sign-change multiplies the left side of (65) by

$$(-1|p_1) \cdot \cdot \cdot (-1|p_s) = (-1|H),$$

and the right side by the same factor; similarly for an interchange of two coördinates, it being noted that $b \equiv b'(2)$.

From $\bar{t}at \equiv mb'(H)$, and hence from (66), follows $[a, mb']_{p_i} = 1$,

$$[a, b']_{p_i} = (m|p_i), \quad \prod [a, b']_{p_i} = (m|H). \tag{67}$$

Similarly $\epsilon(a, mb') = 1$ or -1 according as $Nt = m \equiv 1$ or $-1 \pmod 4$,

$$\epsilon(a, mb') = (-1|m), \quad \epsilon(a, b') = (-h|m)(m|H),$$

the last step by (64'). The theorem will follow when we prove that $(-h|m) = 1$. This is evident from identities (25). For $c = mb'$ and $Nt = m$, whence $x_f^2 + ht_f^2 \equiv 0 \pmod m$ $(f = 0, 1, 2, 3)$. Since some t_f is prime to any prime factor of m, $(-h|m) = 1$.

18. Extension of *AJ* Theorem 7.

THEOREM 10. *Let a and b be proper, pure quaternions of the same norm $h = lH$, $l = 1$ or 2; m an integer prime to $2h$. Then if the g. c. d. of the coördinates of $\bar{t}at$ is prime to h, and Nt and Nu are odd,*

(68) $\qquad \bar{\iota}at \equiv m\bar{u}bu \pmod{4h}$ implies $Nt \equiv (h|m)Nu \pmod 4$.

For, as in (67), since $[a, \bar{\iota}at]_p = 1 = [b, \bar{u}bu]_p$ for each prime factor p of H, and by (65), $\epsilon(a, b) = (m|H)$. Hence by definition of ϵ and (64′),

$$Nt \equiv \epsilon(a, mb)Nu \equiv (h|m)Nu \pmod 4.$$

The case $m = 1$ gives at once

THEOREM 11. *Let h be a positive integer not of the form $4n$ or $8n + 7$. Denote by A the class of all residues $c \pmod{4h}$, such that c_1, c_2, c_3, h are coprime, obtained from $c \equiv \bar{\iota}at \pmod{4h}$ as a ranges over all proper, pure quaternions of norm h, and t over all integer quaternions of norm $\equiv 1 \pmod 4$; by B the class obtained similarly with $Nt \equiv 3 \pmod 4$. The classes A and B are mutually exclusive.*

19. THEOREM 12. *Let $h = lp_1^{r_1} \cdots p_s^{r_s} \not\equiv 7 \pmod 8$, $l = 1$ or 2, the p_i distinct odd primes. Let a be any proper pure quaternion of norm h. For any of the 2^s possible combinations of signs $\epsilon_1, \cdots, \epsilon_s$ (each $\epsilon_i = \pm 1$), we can choose a proper b of norm h such that*

(69) $$[a, b]_{p_i} = \epsilon_i, \qquad (i = 1, \cdots, s).$$

That is, there are pure quaternions of norm h in every possible combination of classes K, L for the various primes dividing $H = h/l$.

Proof. We seek a prime p satisfying

(70) $$(-h|p) = 1, \quad (p|p_i) = \epsilon_i, \qquad (i = 1, \cdots, s).$$

Now $\quad (-h|p) = (-l|p)(H|p) = (-l|p)(p|H)(-1)^{\frac{1}{2}(p-1)\frac{1}{2}(H-1)}$
$\qquad = (-l|p)(-1)^{\frac{1}{2}(p-1)\frac{1}{2}(H-1)}\epsilon_1^{r_1} \cdots \epsilon_s^{r_s}$,

by the latter part of (70). Hence (70_1) can be replaced by a condition on p (mod 4 or 8), and the system (70) satisfied unless $l = 1$ and $H \equiv 3 \pmod 8$, in which case (70) can be satisfied if and only if $\epsilon_1^{r_1} \cdots \epsilon_s^{r_s} = 1$.

In this exceptional case there are an odd number of powers $p_i^{r_i}$ with $p_i \equiv 3 \pmod 4$ and r_i odd. Suppose that 2^{s-1} b's have been chosen, satisfying (69) for every combination of ϵ_i such that $\epsilon_1^{r_1} \cdots \epsilon_s^{r_s} = 1$. Then changing the sign of b_1 changes each of the ϵ_i corresponding to $p_i \equiv 3$ (4) to its negative, and leaves the remaining ϵ_i unchanged (Theorem 3). Clearly we get $2^{s-1}b$'s such that $\epsilon_1^{r_1} \cdots \epsilon_s^{r_s} = -1$, no two having the same set of values $(\epsilon_1, \cdots, \epsilon_s)$.

Having chosen p to satisfy (70) we can by (70_1), choose a_0 so that

$a_0^2 + h \equiv 0 \pmod{p}$; hence the quaternion $a_0 + a$ has a right-divisor of norm p (AJ, p. 896). Say, $a_0 + a = ut$, $Nt = p$. Then $t(a_0 + a)\bar{t} = tu \cdot p$, $ta\bar{t} = pb$, where $b = tu - a_0$. Thus $[a, b]_{p_i} = (p \mid p_i) = \epsilon_i$.

20. The classes A, B of Theorem 11 are exhaustive. More precisely,

THEOREM 13. *Let $h = lH$ ($l = 1$ or 2) be a positive integer not of the form $4n$ or $8n + 7$. As a ranges over all proper, pure quaternions of norm h, and t over all integral quaternions of odd norms, then $\bar{t}at$ represents* (mod $4h$) *all pure quaternion residues c such that*

$$Nc \equiv h \pmod{8h}. \tag{71}$$

For by (71), $Nc \equiv 0 \pmod{H}$. By Theorem 7 (or 6),

$$\bar{t}at \equiv c \pmod{H} \tag{72}$$

is solvable for t if and only if $[a, c]_p = 1$ for each prime p dividing H (which can be satisfied, in view of Theorem 12, by choice of a) and the norm-condition $s^2 Na \equiv Nc \pmod{H}$ can be solved. Since $H \mid Na$ and Nc, s is arbitrary.

Further, by (71), $Na \equiv Nc \pmod{8l}$ for any a of norm h, and hence by an *even* number of permutations of the coördinates of a (thereby not affecting the solvability of (72)), we can secure $a \equiv c$ (2). By AJ, Theorem 4,

$$\bar{t}at \equiv c \pmod{4l} \tag{73}$$

is solvable. The moduli being coprime both (72) and (73) can be satisfied. From Corollary 3 follows immediately the

COROLLARY. *Theorem 13 holds with Nt confined to any residue class* (mod H).

21. Extension of *AJ* Theorem 1.

THEOREM 14. *If $8n + 1$ is not a square, and h is a positive integer not of the form $4k$ or $8k + 7$, the equation* (1′) *has equally many proper solutions in each of the classes A and B defined in Theorem 11. If $8n + 1 = m^2$ $(m > 0)$, then all proper solutions are in A or B according as $m \equiv 1$ or 3* (mod 4).

By AJ, p. 896, any proper pure quaternion of norm hm^2 is of the form $\bar{t}at$, with $Na = h$ and $Nt = m$. Hence the second part follows from Theorem 11.

Assume $8n + 1$ not square. We shall set up a $(4 - 4)$ correspondence between the solutions of types A and B. For this we choose a prime p such that

$$(74) \qquad (-h|p) = -1, \quad (-h(8n+1)|p) = 1.$$

This is equivalent to the following:

$$(-l|p)(-1)^{\frac{1}{2}(p-1)\cdot\frac{1}{2}(H-1)}(p|H) = -1, \quad (p|8n+1) = -1,$$

and can evidently be satisfied except possibly when $l=1$ and $H \equiv 3\ (8)$; it then reduces to $(p|H) = -1$, $(p|8n+1) = -1$; these are still evidently compatible unless H and $8n+1$ contain exactly the same prime factors, say

$$H = p_1^{e_1}p_2^{e_2}\cdots p_s^{e_s} \quad \text{and} \quad 8n+1 = p_1^{f_1}p_2^{f_2}\cdots p_s^{f_s}.$$

At least one e_i is odd since $H \equiv 3\ (8)$, and at least one f_i since $8n+1$ is not a square. If e_i and f_i are odd for the same i, we choose $(p|p_i) = -1$, and $(p|p_k) = 1$ for $k \neq i$; otherwise, for some $i \neq j$, e_i and f_j are odd, e_j and f_i even; and $(p|p_i) = (p|p_j) = -1$, the remainder $+1$, is effective.

Choose a solution x_0 of $x_0^2 + h(8n+1) \equiv 0\ (p)$. We employ the following process with p and x_0 if x is a proper solution of (1′) of class A, with p and $-x_0$ if x is of class B. Since $p|N(x_0+x)$,

$$(75) \qquad x_0 + x = uv, \quad Nv = p,$$

v being unique up to a left unit factor. Hence $vx\bar{v} = px'$, where $x' = vu - x_0$. Taking conjugates gives

$$(76) \qquad -x_0 + x' = (-\bar{u})\bar{v}, \text{ where } N\bar{v} = p.$$

Hence if the process be applied to x', with p and $-x_0$, the right-divisors obtained are the left-associates of $\bar{v}$, and since $\bar{v}x'v = px$, we are led back to x. If $i_a v$ is employed in place of v in (75), x' becomes $-i_a x' i_a$ which differs from x' in the signs of two of x'_1, x'_2, x'_3 only. Thus the four quaternions

$$(77) \quad (x_1, x_2, x_3), \quad (x_1, -x_2, -x_3), \quad (-x_1, x_2, -x_3), \quad (-x_1, -x_2, x_3)$$

are associated with four similarly related quaternions x'. The $(4-4)$ correspondence will be established when we show that if x is of class A, x' is of class B, and conversely; for the method of choosing between x_0 and $-x_0$ assures that no x not already placed in the correspondence will lead to one already placed. By Theorem 13 we can set $x \equiv \bar{t}at$, $x' \equiv \bar{w}aw \pmod{4h}$; thus $v\bar{t}at\bar{v} \equiv p\bar{w}aw\ (4h)$; by Theorem 10, since $Nv \equiv (-1|p) \pmod 4$, $Nt \equiv (-h|p)Nw \equiv -Nw\ (4)$, the last step by (74_1).

22. Theorem 15. *For a and h as in Theorem 12 there are equally*

many proper, pure quaternions b of norm h having any of the 2^s sets of values $\epsilon_1, \cdots, \epsilon_s$.

In the proof of Theorem 12 we chose primes $p = p(\epsilon_1, \cdots, \epsilon_s)$ and with each p an $a_0 = a_0(p)$. We use a_0 when applying the process to a quaternion having every $\epsilon_i = 1$, otherwise $-a_0$. Then much as in the preceding section, with every quadruplet (77) having every $\epsilon_i = 1$ is associated 2^s or 2^{s-1} quadruplets b with the respective values $\epsilon_1, \cdots, \epsilon_s$; the 2^{s-1} applying only if $h \equiv 3\ (8)$, when $\epsilon_1^{r_1} \cdots \epsilon_s^{r_s} = 1$. If the process, with the corresponding p and $-a_0$, be applied to any of these b's we are led back to the original four x's. If $\epsilon_1 = \cdots = \epsilon_s = 1$, we take $p = 1$, so that (77) then leads to itself.

If a b distinct from those already considered exists, there is one with every $\epsilon_i = 1$. For applying to such a b the corresponding $p(\epsilon_1, \cdots, \epsilon_s)$ and $-a_0(p)$, we obtain one with every $\epsilon_i = 1$ and distinct from those already considered, since b is. The theorem follows, but only for $\epsilon_1^{r_1} \cdots, \epsilon_s^{r_s} = 1$ if $h \equiv 3\ (8)$. Then, changing the sign of b_1 evidently produces an equal number of b's for each $(\epsilon_1, \cdots, \epsilon_s)$ such that $\epsilon_1^{r_1} \cdots \epsilon_s^{r_s} = -1$.

Let $R_3(h)$ denote the number of proper representations of h as a sum of three squares. In view of Theorem 3 we have the

COROLLARY. *Let s denote the number of distinct odd prime factors of h, $h > 3$, $h \neq 4n$, $h \neq 8n + 7$. Then $4 \cdot 2^s \mid R_3(h)$, and $R_3(h) \geqq 12 \cdot 2^s$.*

For in case of no equalities or zeros, the 48 proper representations obtained from (x_1, x_2, x_3) by permutations and sign-changes have at most four values $(\epsilon_1, \cdots, \epsilon_s)$; a proper representation $(x_1, x_2, 0)$ or (x_1, x_2, x_2) can occur only if $(-1 \mid p_i) = 1$ (or $(-2 \mid p_i) = 1$) for every p_i, and then there are at most two values $(\epsilon_1, \cdots, \epsilon_s)$.

WESTON COLLEGE
AND
McGILL UNIVERSITY.

QUASI-GROUPS WHICH SATISFY CERTAIN GENERALIZED ASSOCIATIVE LAWS.*

By D. C. Murdoch.[1]

Introduction. In a recent paper[2] on the theory of quasi-groups, B. A. Hausmann and O. Ore have given an interesting analysis of the associative law, and have shown that it may be generalized to a very considerable extent without invalidating many of the principal theorems of group theory. In the present paper we shall consider an interesting class of quasi-groups which are not for the most part included under the laws discussed by Hausmann and Ore, namely, those which satisfy an associative law of the form

$$a(bc) = (ab)c_1,$$

where c_1 is independent of b.

In the first section properties of the right units of such a quasi-group G are derived. It is shown that G contains a set of minimal right unit sub-quasi-groups having no elements in common. At least one of these is contained in every sub-quasi-group of G.

The second section deals with coset expansions. In order to obtain a suitable definition of a normal sub-quasi-group, it is necessary to assume, in addition, a second associative law, symmetrical to the first, namely,

$$a(bc) = (a_1b)c$$

where a_1 is independent of b. Certain structural properties of such quasi-groups, and the Jordan-Hölder theorem, are proved in Section 3.

The fourth section is devoted to quasi-groups which satisfy the law

$$(ab)(cd) = (ac)(bd).$$

This law is interesting since it not only implies both previous laws, but at the same time imposes a generalized commutative law. In fact, such quasi-groups are direct generalizations of Abelian groups. For this reason I have called them Abelian quasi-groups, in spite of the fact that they are not in general commutative.

---

* Received May 19, 1938; revised October 15, 1938.

[1] Sterling Fellow, Yale University.

[2] B. A. Hausmann and Oystein Ore, "Theory of quasi-groups," *American Journal of Mathematics*, vol. 59 (1937), pp. 983-1004.

In Section 5 are discussed properties of quasi-groups which satisfy Law 1 and have a unique right unit. These are the same as those satisfying postulate B of A. Suschkewitsch,[3] and have properties very similar to those of ordinary groups. The final section is devoted to examples, and to methods of constructing various types of quasi-groups.

1. Quasi-groups which satisfy Law 1. Right Units. By a quasi-group G we shall understand a set of elements, closed under multiplication, in which the quotient axiom is satisfied. That is, the equations $ax = b$, $ya = b$ are uniquely soluble for x and y where a and b are any two (not necessarily distinct) elements of G. This implies both left and right cancellation laws. We shall be concerned only with finite quasi-groups, although many of the results obtained hold in the infinite case also. If G is finite, then every subset of G, which is closed under multiplication, satisfies the quotient axiom, and is therefore a sub-quasi-group of G.

We shall assume that G satisfies:

Associative Law I. *If a, b, c are any three elements of G then*

$$a(bc) = (ab)c_1, \tag{1}$$

where c_1 is independent of b.

From the quotient axiom it follows that every element a of G has a right unit e_a and a left unit e'_a, defined by the equations

$$ae_a = e'_a a = a.$$

These units are uniquely determined. Since in (1), c_1 is independent of b, we find on putting $b = e_a$ that

$$a(e_a c) = ac_1,$$

and therefore $c_1 = e_a c$. It is convenient now to introduce the following notation: denote by $f_a(c)$ the element defined by the equation

$$e_a f_a(c) = c,$$

where c is any element of G and a any fixed element. It follows that

$$f_a(e_a c) = c,$$

and hence the function f_a^{-1} inverse to f_a is defined by

[3] A. Suschkewitsch, "On a generalization of the associative law," *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 204-14.

$$f_a^{-1}(c) = e_a c.$$

Our associative law (1) then becomes

$$(2) \qquad \begin{aligned} a(bc) &= (ab)f_a^{-1}(c), \\ (ab)c &= a[bf_a(c)], \end{aligned}$$

where the second equation is a consequence of the first.

If we put $c = e_b$ in the first equation of (2), we find that

$$e_{ab} = e_a e_b,$$

from which follows:

THEOREM 1. *The set of all right units of G form a sub-quasi-group R, and $a \rightarrow e_a$ is a homomorphic mapping of G on R.*

The homomorphism $a \rightarrow e_a$ maps G on R, R on its right unit quasi-group R_1, R_1 on its right unit quasi-group R_2, and so on. Since G is finite we must finally reach a sub-quasi-group R_t which is mapped on itself. Since every sub-quasi-group of R_t must contain its own right units, it follows that $a \rightarrow e_a$ is an automorphism, not only of R_t but also of every sub-quasi-group of R_t. Now let $E, E_2, \cdots, E_r$ be the set of all minimal sub-quasi-groups of R_t. Since every sub-quasi-group of G contains a sub-quasi-group of R, and therefore of $R_1, R_2, \cdots, R_t$, it must contain at least one of these E_i. These results may be stated as follows:

THEOREM 2. *Any finite quasi-group G satisfying Law 1 contains a set of minimal right unit sub-quasi-groups, no two of which have elements in common, and at least one of which is contained in every sub-quasi-group of G.*

If one of the minimal right unit sub-quasi-groups consists of a single element e, then e is its own right (and left) unit and will be called a *principal unit.* The set of all principal units of G form a sub-quasi-group, since if $e_1 e_1 = e_1$ and $e_2 e_2 = e_2$,

$$e_1 e_2 = e_1(e_2 e_2) = (e_1 e_2)(e_1 e_2),$$

and hence $e_1 e_2$ is also a principal unit.

Since the minimal right unit sub-quasi-groups themselves have no sub-quasi-groups, they are necessarily cyclic quasi-groups in which every element is primitive. Moreover, every element is the right unit of some other element. They are not necessarily of prime order as might be expected, since the direct product of two distinct quasi-groups which have no sub-quasi-groups, will itself have no sub-quasi-groups.

2. Coset expansions and normal sub-quasi-groups. A necessary and sufficient condition for the existence of left coset expansions with respect to any sub-quasi-group is obtained by Hausmann and Ore and may be stated as follows:

If, for arbitrary elements a and b and a fixed element c_0,

$$(ab)c_0 = ad_0,$$

then for arbitrary c

$$(ab)c = ad,$$

where d belongs to the quasi-group $\{c_0, d_0, c\}$.

This condition is easily seen to be satisfied in our case if we put $c_0 = e_a$. For then $d_0 = bf_a(e_a)$, and hence b, and therefore $bf_a(c)$ belongs to $\{e_a, d_0, c\}$. Hence we have

THEOREM 3. *If H is any sub-quasi-group of G, then G may be represented by means of left cosets of H.*

If a sub-quasi-group H does not contain the right unit quasi-group R of G, then a coset aH will not always contain its defining element a. It is natural to ask whether all elements of the coset containing a define the same coset aH. This is easily proved to be so if we make G symmetrical by assuming, in addition to (1),

ASSOCIATIVE LAW II. *If a, b, c are any three elements of G, then*

$$(ab)c = a_1(bc), \tag{3}$$

where a_1 is independent of b.

Since a_1 is independent of b, we find on putting $b = e'_c$, that

$$(ae'_c)c = a_1c,$$

and therefore $a_1 = ae'_c$. Now if we define the function f'_c by the equation

$$f'_c(a)e'_c = a,$$

for all a and any fixed c in G, we find that

$$f'_c(ae'_c) = a,$$

and therefore the inverse function f'^{-1}_c is defined by

$$f'^{-1}_c(a) = ae'_c.$$

Hence our second associative law (3) becomes

$$(4) \qquad \begin{aligned} (ab)c &= f'^{-1}_c(a)(bc), \\ a(bc) &= [f'_c(a)b]c, \end{aligned}$$

the second equation being a consequence of the first.

From (2) and (4) we have now

$$a[bf_a(c)] = (ae'_c)(bc) = a[e'_c f_a(bc)],$$

and therefore

$$(5) \qquad bf_a(c) = e'_c f_a(bc),$$

for all elements a, b, and c in G. Now if we let c run through all elements of a sub-quasi-group H, while b is any fixed element of H, say h, the right hand side of (3) gives the same set of elements whatever element h is chosen. Hence the left hand side, $hf_a(H)$ is the same for any element h of H. The same is therefore true of

$$(ah)H = a[hf_a(H)]$$

and we have proved

THEOREM 4. *If G satisfies Associative Laws I and II, and H is any sub-quasi-group of G, then all elements of a coset aH define a fixed coset $(ah)H$ which is equal to aH if and only if H contains e_a.*

It also follows from (4), on putting $a = e'_b$, that the left units of G form a sub-quasi-group L, and $a \to e'_a$ is a homomorphism mapping G on L. By an argument similar to that of Theorem 2, we obtain a set of minimal left unit sub-quasi-groups $E'_1, E'_2, \cdots, E'_s$, one of which must be contained in every sub-quasi-group of G. It follows that these are identical, in some order, with the minimal right unit sub-quasi-groups, and may therefore be called, without ambiguity, minimal unit sub-quasi-groups. Law II also insures the existence of right coset expansions.

A further consequence of equation (5) is:

THEOREM 5. *The product set $(aH)(bH)$ of any two cosets contains the coset $(ab)H$.*

For if c runs through all elements of H, (5) gives

$$(6) \qquad bf_a(H) \subseteq Hf_a(bH)$$

for all a and b in G, and therefore

$$(ab)H = a[bf_a(H)] \subseteq a[Hf_a(bH)] = (aH)(bH).$$

If equality holds in (6), that is if

$$Hf_a(bH) = bf_a(H)$$

for all a and b in G, then H will be called a *left-normal* sub-quasi-group of G. If H is left normal in G then

$$(aH)(bH) = (ab)H,$$

and the cosets of H evidently form a left quotient quasi-group G/H which is homomorphic to G.

THEOREM 6. *If H is left normal in G, the mapping of each coset aH on the coset defined by any element of aH is an automorphism of G/H.*

For

$$[(ah)H][(bh)H] = [(ah)(bh)]H = [(ab)h]H.$$

Evidently right normality could be defined by

$$f'_b(Ha)H = f'_b(H)a$$

and a right quotient group would result. Since all theorems which hold in the one case will hold in the other also, we shall confine out attention to the left hand case. Unless otherwise stated, we shall use the terms coset, normal, and quotient quasi-group to mean left coset, left normal and left quotient quasi-group respectively.

3. Structural properties and the Jordan-Hölder Theorem. It is evident that if we include a void sub-quasi-group O, containing no elements, and contained in every sub-quasi-group of G, then the sub-quasi-groups of G will form a structure, or lattice. In order to prove the same for the normal sub-quasi-groups, it is necessary to confine ourselves to those normal sub-quasi-groups which contain the right unit quasi-group R of G. We shall show that these not only form a structure, but a Dedekind structure.

Two sub-quasi-groups H and K of G will be said to be *permutable* if for any two elements h and k we have

$$hk = k'h',$$

where k' is in K and h' in H.

THEOREM 7. *Any two normal sub-quasi-groups H and K which have a non-void crosscut D, are permutable.*

Proof. If h, k are any elements of H, K respectively, and d is any fixed element of D, then

$$hk = (dh_1)(k_1d) = d[h_1f_d(k_1d)] = d[k_1f_d(h_2)] = (dk_1)h_2 = k_2h_2,$$

where k_2 is in K and h_2 is in H. Similarly from the normality of K, every product kh may be written $h'k'$.

COROLLARY. *If the crosscut* (H,K) *is not void, and if* H *and* K *are normal, then the union* $[H,K]$ *consists of all elements, and only those, of the form* hk.

THEOREM 8. *The normal sub-quasi-groups of* G *which contain* R *form a Dedekind structure.*

Proof. To show that they form a structure, it is sufficient to show that the union of two such sub-quasi-groups is normal, since this is obvious for the crosscut. For a sub-quasi-group H containing R the normality condition reduces to

$$H(bH) = f_a^{-1}(b)H \tag{7}$$

for all elements a and b in G. By the above corollary we have $[H,K] = HK$ and from Theorem 7 and the normality of H and K,

$$\begin{aligned}(HK)[b(HK)] &= (KH)[(bH)K] = [(KH)(bH)]K \\ &= [(Kb)H]K = K[(bH)K] = f_a^{-1}(bH)K = f_a^{-1}(b)[HK]\end{aligned}$$

for all a and b in G. Hence $[H,K]$ is normal, and the normal sub-quasi-groups containing R form a structure, Σ.

To show that it is a Dedekind structure, it is necessary to show that if M is any element of Σ containing H, then

$$(M, [H, K]) = [H, (M, K)].$$

The right hand side is certainly contained in the left, and that the left is also contained in the right follows easily from the corollary to Theorem 7.[4]

The Dedekind structure Σ will contain a unit element $\bar{R}$ which will be equal to R if R is normal. It follows from the properties of Dedekind structures[5] that all principal chains of normal sub-quasi-groups between G and $\bar{R}$ have the same length, and that the quotient structures between successive terms in any one such principal chain are isomorphic in some order to those in any other.

We can also prove the law of isomorphism just as in the case of groups:

---

[4] See Hausmann and Ore, page 995, Theorem 4.

[5] Oystein Ore, "On the foundation of abstract algebra I," *Annals of Mathematics*, vol. 36 (1935), no. 2.

THEOREM 9. *If H and K are any two permutable sub-quasi-groups of G which both contain R, and if H is normal in the union $[H, K]$, then the crosscut (H, K) is normal in K and*

$$[H, K]/H \cong K/(H, K).$$

The Jordan-Hölder theorem then follows, as in groups, for series of composition between G and $\bar{R}$. These theorems, of course, give us no information concerning quasi-groups in which $G = R$, or in which $G = \bar{R} \neq R$. There are two special cases, however, which are of considerable interest and concerning which more can be said. These are the 'Abelian' quasi-groups of the next section, and those in which R consists of a single element, which will be treated in Section 5. The latter restriction is a great simplification and leads to quasi-groups which are a special case of those discussed by Suschkewitsch, and which preserve many properties of ordinary groups.

4. Abelian quasi-groups. In this section we shall consider quasi-groups which satisfy

ASSOCIATIVE LAW III. *If a, b, c, d are any four elements of G, then*

$$(ab)(cd) = (ac)(bd). \tag{8}$$

Such a quasi-group will be called *Abelian* since, although not in general commutative, it is, as we shall see, a generalization of an Abelian group.

If a is any element of G we shall understand by a power of a, any element of the cyclic quasi-group generated by a. Such a power will be represented by the notation $\phi_r(a)$ where r is the number of factors a which occur. We can then prove the following theorem which generalizes the law $(ab)^r = a^r b^r$ of Abelian groups.

THEOREM 10. *If a and b are any two elements of an Abelian quasi-group, and if $\phi_n(a)$ is any power of a, then*

$$\phi_n(ab) = \phi_n(a)\phi_n(b). \tag{9}$$

Proof. From (8) the theorem obviously holds for $n = 2$. Assuming it true for all powers involving fewer than n factors, since every power $\phi_n(a)$ can be written as the product of two such powers, we have, for some r less than n,

$$\begin{aligned} \phi_n(ab) &= \psi_r(ab)\chi_{n-r}(ab) = [\psi_r(a)\psi_r(b)][\chi_{n-r}(a)\chi_{n-r}(b)] \\ &= [\psi_r(a)\chi_{n-r}(a)][\psi_r(b)\chi_{n-r}(b)] = \phi_n(a)\phi_n(b). \end{aligned}$$

The theorem therefore holds for powers with n factors, and therefore holds in general.

COROLLARY. *If $\phi_n(a)$, $\psi_m(a)$ are any two powers of a then*

$$\phi_n[\psi_m(a)] = \psi_m[\phi_n(a)].$$

This follows by repeated application of the formula which results from (9) on putting $b = a$.

It is evident that in an Abelian quasi-group all sub-quasi-groups are normal. We therefore have $\bar{R} = R$, and Theorems 8 and 9 will give the Jordan-Hölder theorem for composition series (which are here the same as the principal series of normal sub-quasi-groups) between G and R. Moreover, since the union of any two normal sub-quasi-groups is now necessarily normal, we can prove, just as in Theorem 8,

THEOREM 11. *In an Abelian quasi-group, the set of all sub-quasi-groups which contain a given minimal unit sub-quasi-group, forms a Dedekind structure.*

Hence all principal chains between G and any fixed minimal unit sub-quasi-group E_i must have the same length, and the quotient quasi-groups associated with any two of these chains will be isomorphic in some order. However, the set of all sub-quasi-groups (including a void one) will not in general form a Dedekind structure, and therefore principal chains between G and two different minimal unit sub-quasi-groups, need not have the same length.

If $H \supseteq R$, the quotient quasi-group G/H will have a unique right unit, since $(aH)H = a[Hf_a(H)] = aH$, for all a in G. In particular this is true of R itself, and since G/R satisfies the same associative law as G, we find that left multiplication by R is an automorphism s of G/R and, in view of (7), we may write

$$R(cR) = (cR)^s = f_a^{-1}(c)R$$

for all a and c in G. Putting $a = e_a$, $d = e_c$ in (8), we find

$$f_a^{-1}(b)c = f_a^{-1}(c)(be_c).$$

Now let b run through all elements of R, and we have

$$Rc = f_a^{-1}(c)R,$$

and therefore,

$$R(cR) = (cR)^s = Rc.$$

These results may be stated as follows:

THEOREM 12. *If G is Abelian, the quotient quasi-group G/R has a unique*

right unit R. Every left coset of R is also a right coset of R and conversely. The mapping $aR \rightarrow Ra$ is an automorphism of G/R and is equivalent to left multiplication by the right unit R. Finally if $aR = Ra$ for all a in G then G/R is a group.

This last statement follows since if a quasi-group satisfying any one of our associative laws has an absolute unit, it is a group.

5. Quasi-groups with unique right unit. If G satisfies Law I and has a unique right unit e, then equations (2) become

$$\begin{aligned} a(bc) &= (ab)c^s \\ (ab)c &= a(bc^{s^{-1}}) \end{aligned} \tag{11}$$

where $c^s = ec$ and $ec^{s^{-1}} = c$. Putting $a = e$ in the first equation of (11), we find

$$(bc)^s = b^s c^s$$

and therefore s is an automorphism of G.

Right and left inverses of any element a are uniquely defined by

$$aa^{-1} = a_{-1}a = e,$$

and the solutions of the linear equations $xa = b$, $ay = b$ are then found to be

$$x = b(a^{-1})_s\, b(a^{-1})^s, \quad y = (a_{-1}b)^{s^{-2}}. \tag{13}$$

Hence, putting $b = e$

$$a_{-1} = (a^{-1})^{s^2}. \tag{14}$$

Evidently

$$(a^{-1})_{-1} = (a_{-1})^{-1} = a,$$

and therefore

$$(a^{-1})^{-1} = a^{s^{-2}}, \quad (a_{-1})_{-1} = a^{s^2}.$$

THEOREM 13. *The set H, of all elements which commute with e, is a group, the largest group contained in G.*

This follows from (12), (14), and (11), since $a^s = a$ for all elements a of H.

Since e is contained in every sub-quasi-group H of G, every left coset aH contains its defining element a. In this case, therefore, it is not necessary

to assume Law II in order to prove Theorems 4 and 5. Both these Theorems are obviously true whenever H contains the right unit quasi-group R. The definition of normality also becomes greatly simplified. A normal sub-quasi-group H, may be defined as one for which

$$a^s H = Ha$$

for every element a of G. Then if H is normal, we have

$$H(aH) = a^s H = Ha$$

and therefore $aH \rightarrow Ha$ is an automorphism of G/H, and is equivalent to left multiplication by the right unit. If $aH = Ha$ for all a in G, then G/H is a group. This is the case only if H contains the left units of all elements of G. Since $R = \bar{R} = e$, the normal sub-quasi-groups form a Dedekind structure and the law of isomorphism and the Jordan-Hölder theorem hold exactly as in groups.

If G is Abelian, putting $a = d = e$ in (8), we find that

$$b^s c = c^s b \tag{15}$$

for all b and c in G. Hence when G has a unique right unit, (8) implies both (11) and (15). Conversely (11) and (15) imply (8), for

$$(ab)(cd) = a[b(cd)^{s^{-1}}] = a[(b^s c)d^s]^{s^{-1}} = a[(c^s b)d^s]^{s^{-1}} = (ac)(bd).$$

Therefore, in the case of quasi-groups with unique right unit satisfying Law I, the Abelian quasi-groups may be characterized by (15).

Let G be any Abelian quasi-group with unique right unit e, and let $\phi(a)$ denote any power of a. It follows from Theorem 10, that to any such power ϕ there correspond two sub-quasi-groups of G. The first, which we shall denote by G_ϕ, consists of all elements x of G which satisfy the relation

$$\phi(x) = e,$$

and the second, $G^{(\phi)}$, consists of all elements of the form $\phi(x)$, where x runs through all elements of G. We can then prove

THEOREM 14. *The quotient quasi-group G/G_ϕ is isomorphic to $G^{(\phi)}$.*

Proof. $\phi(a) = \phi(b)$ if and only if b lies in aG_ϕ. For if $b = ag$ where g is an element of G_ϕ, then $\phi(b) = \phi(ag) = \phi(a)\phi(g) = \phi(a)$. Conversely if $\phi(a) = \phi(b)$, then

$$\phi(a_{-1}b) = \phi(a_{-1})\phi(b) = \phi(a_{-1})\phi(a) = \phi(a_{-1}a) = e,$$

and therefore $a_{-1}b$ belongs to G_ϕ, and by (13) and subsequent equations, b belongs to aG_ϕ. Hence the correspondence

$$aG_\phi \leftrightarrow \phi(a)$$

is evidently an isomorphism between G/G_ϕ and $G^{(\phi)}$.

The transform a_b, of a by b, may be defined by the equation

$$(16) \qquad a_b = (ba^{s^{-1}})b^{-1},$$

and the laws of transformation are easily deduced. Moreover the set C of all elements which are invariant under transformation by all elements of G, forms an Abelian sub-quasi-group, the centre of G.

Quasi-groups of this type are a special case of those discussed by Suschkewitsch, who showed that they always contain a second operation $\times$, under which they form a group. In our case this operation is defined by

$$b \times c = bc^{s^{-1}}$$

and is obviously associative. Equation (16) then becomes

$$a_b = b \times a \times b^{-1^s}$$

which is the ordinary transform of a by b in the group $G(\times)$, since b^{-1^s} is the inverse of b with respect to the operation $\times$.

Although every sub-quasi-group H of G gives rise to a sub-group $H(\times)$ of $G(\times)$, the converse[6] is not necessarily true. For example, the *normalizer* of any set of elements H of G can be defined as the set N of all elements a of G such that

$$(aH^{s^{-1}})a^{-1} = H.$$

If a and b are in N, it follows from the laws of transformation that $ab^{s^{-1}}$ is also in N and therefore $N(\times)$ is a sub-group of $G(\times)$, and the order of N divides that of G, but we cannot say that N is necessarily a sub-quasi-group of G. It can be shown, however, just as in groups, that the number of distinct sets conjugate to H is g/n, where g is the order of G and n the order of N.

6. Examples. The quasi-groups with unique right unit, which satisfy (11), are identical with those satisfying Postulate B of A. Suschkewitsch.

---

[6] Suschkewitsch shows that this converse holds if and only if G is obtained from an ordinary group by making a substitution $\begin{pmatrix} X \\ X^r \end{pmatrix}$ in the headline of the group table, where the exponents r are relatively prime to the orders of the corresponding elements X.

They are the only ones included under both Law I of the present paper and the general law $(ab)c = ad$, where d is independent of a, discussed by Suschkewitsch. We shall refer to them for convenience as quasi-groups of type B.

Suschkewitsch showed that any quasi-group of type B can be obtained from an ordinary group $G(\times)$ by making a permutation s in the horizontal title line of the group table, where s is an automorphism of the group. For example, the set of all 'vectors' of the form

$$a = (a_1, a_2, \cdots, a_n) \tag{17}$$

where $a_1, \cdots, a_n$ are arbitrary elements of a given group $G(\times)$, will form a quasi-group of type B, if we define the product of a by a similar symbol b to be

$$ab = (a_1b_n, a_2b_1, \cdots, a_nb_{n-1}).$$

This merely amounts to making a Suschkewitsch substitution in the direct product of n factors $G(\times)$.

A large class of Abelian quasi-groups can be obtained by a simple extension of this method. Let $G(\times)$ be any Abelian group, in which multiplication is denoted by $\times$, and let s and t be any two automorphisms of $G(\times)$ such that $st = ts$. The elements of $G(\times)$ then form an Abelian quasi-group G, if multiplication in G is defined by

$$ab = a^s \times b^t. \tag{18}$$

For then

$$(ab)(cd) = a^{s^2} \times b^{ts} \times c^{st} \times d^{t^2} = (ac)(bd).$$

Moreover, s and t are also automorphisms of G. If s is the identity, this is equivalent to making a Suschkewitsch substitution t^{-1} in the title line of $G(\times)$, and the resulting quasi-group is Abelian of type B. It easily follows that a quasi-group of type B is Abelian if, and only if, it is obtained by a Suschkewitsch substitution from an Abelian group. The symbols (17) will form an Abelian quasi-group, not of type B, if $a_1, a_2, \cdots, a_n$ are arbitrary elements of an Abelian group $G(\times)$ and if multiplication is defined by

$$ab = (a_2b_n, a_3b_1, \cdots, a_nb_{n-2}, a_1b_{n-1}). \tag{19}$$

This again, amounts to performing a substitution of type (18) on the direct product of n factors $G(\times)$. Equation (18) is, of course, equivalent to performing the substitutions s^{-1} and t^{-1} on the vertical and horizontal title lines

18

respectively, in the group table for $G(\times)$. An Abelian quasi-group formed in this way will be commutative if and only if $s = t$. It follows that if an Abelian quasi-group of type B is commutative, it is an Abelian group.

The above method, however, does not give all Abelian quasi-groups. For every quasi-group obtained in this way will contain at least one principal unit, the unit elemtnt of $G(\times)$. That Abelian quasi-groups exist containing no principal unit is shown by example (3) below. The quasi-groups of order 3,

(1)

	1	2	3
1	1	3	2
2	2	1	3
3	3	2	1

(2)

	1	2	3
1	1	3	2
2	3	2	1
3	2	1	3

(3)

	1	2	3
1	2	1	3
2	1	3	2
3	3	2	1,

illustrate three different types. They are all Abelian; (1) has a unique right unit; in (2) every element is a principal unit; (3) contains no principal unit and has no sub-quasi-groups. The direct product of (3) with another quasi-group having no sub-quasi-groups would give a quasi-group of composite order, but with no sub-quasi-groups. Sylow's theorem, therefore, does not hold in general. Also (2) illustrates the fact that an Abelian quasi-group is not necessarily the direct product of cyclic quasi-groups. A non-Abelian cyclic quasi-group of type B, is given at the end of Suschkewitsch's paper. It is obtained from the symmetric group of order 6.

Of interest, too, is the following Abelian quasi-group of order 9 which cannot be obtained directly by the method described above:

	1	2	3	4	5	6	7	8	9
1	1	2	3	5	6	4	9	7	8
2	2	3	1	6	4	5	7	8	9
3	3	1	2	4	5	6	8	9	7
4	6	4	5	7	8	9	2	3	1
5	4	5	6	8	9	7	3	1	2
6	5	6	4	9	7	8	1	2	3
7	8	9	7	3	1	2	4	5	6
8	9	7	8	1	2	3	5	6	4
9	7	8	9	2	3	1	6	4	5

It is the right unit quasi-group of the quasi-group of order 27 consisting of all elements of the form (17) where $n = 3$, a_1, a_2, a_3 are arbitrary elements of the cyclic group of order 3, and multiplication is defined by (19).

YALE UNIVERSITY.

ON DIFFERENTIAL OPERATORS IN HILBERT SPACES.*

By Kurt Friedrichs.

Symmetric differential operators from the point of view of Hilbert space present principally two problems.

The first problem is to define the domain of the operator (i. e., the manifold of functions on which it is applicable) in such a way that the operator is self-adjoint (or hypermaximal). According to the theory of v. Neumann and Stone the problem of obtaining the spectral resolution of a non-bounded operator is completely reduced to that of establishing its self-adjointness.[1]

The second type of problems refers to the properties of this domain,[2] in particular the question arises under which circumstances functions of this domain are continuous and have derivatives.

The first problem may be exemplified in the case of the operator

$$-\Delta = -\frac{\partial}{\partial x_1{}^2} - \cdots - \frac{\partial}{\partial x_n{}^2}$$

acting on functions $u(x_1, \cdots, x_n)$ which are defined in a region S and vanish at the boundary of S. A theorem of v. Neumann [3] states that an operator is self-adjoint if it is the product of two adjoint operators. This suggests considering the above operator as the product D^*D of the gradient D, which transforms the function u into the system

$$Du = \left\{ \frac{\partial u}{\partial x_1}, \cdots, \frac{\partial u}{\partial x_n} \right\}$$

and the negative divergence D^*, which transforms a system $v' = \{v_1, \cdots, v_n\}$ into the function

$$D^*v' = -\frac{\partial v_1}{\partial x_1} - \cdots - \frac{\partial v_n}{\partial x_n}.$$

---

* Received May 25, 1938.

[1] It may be noted that, on the basis of this paper, the theory of the corresponding boundary value and eigenvalue problems is, therefore, obtained by straightforward application of the existing theory of the Hilbert space. For a different approach cf. Courant-Hilbert, *Methoden der mathematischen Physik*, vol. II, chap. VII.

[2] Cf. I. Halperin, "Closures and adjoints of linear differential operators," *Annals of Mathematics*, vol. 38 (1937), p. 880. A complete characterization of closure and adjoints is given for ordinary and hyperbolic differential operators of any order and certain elliptic cases are treated. Cf. also F. J. Murray, *Transactions of the American Mathematical Society*, vol. 37 (1935), p. 301.

[3] J. v. Neumann, *Annals of Mathematics*, vol. 33 (1932), p. 294.

Thus the problem is reduced to that of defining the domains ($\mathfrak{G}_\infty$ and $\mathfrak{G}'$) of D and D^* so that these operators are adjoint. The domain $\mathfrak{G}_\infty$ of D is restricted by the conditions

$$\int \cdots \int_S u^2 dx_1 \cdots dx_n < \infty,$$

$$\int \cdots \int_S \left\{ \left(\frac{\partial u}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial u}{\partial x_n}\right)^2 \right\} dx_1 \cdots dx_n < \infty,$$

and by the boundary condition; the domain G' of D^* is restricted by

$$\int \cdots \int_S \{v_1 + \cdots + v_n^2\}\, dx_1 \cdots dx_n < \infty$$

and

$$\int \cdots \int_S (D^*v')^2 dx_1 \cdots dx_n < \infty.$$

In order that the operators D and D^* are closed the class of functions u and $v_1, \cdots, v_n$ with continuous derivatives must be extended to manifolds of functions which are differentiable in a more general sense. For the operator D, this extension is possible by permitting differentiability in the sense of Lebesgue's theory. A correspondent definition of D^* does not seem to be obvious. Therefore, it seems preferable to effect the extension of both operators in the following different way giving no preference to D or D^*.

First, we consider the operator D, applied on functions which have continuous derivatives and vanish in the neighborhood of the boundary of S. We then define D in $\mathfrak{G}_\infty$ as the closure,[4] and D^* in $\mathfrak{G}'$ as the adjoint of this operator. We say: D in $\mathfrak{G}_\infty$ is defined in the strong, and D^* in $\mathfrak{G}'$ in the weak sense.

Secondly, we consider the operator D^*, applied on systems of functions with continuous derivatives and define D^* in $\mathfrak{G}'$ as the closure (strong sense), and $\mathfrak{D}$ in G_∞ as the adjoint (weak sense) of this operator.

Our main theorem is that the strong and the weak extensions coincide.[5]

To prove this identity we use a sequence of a simple integral operators, which produce functions with derivatives of every order, which are commutative with D and D^* even in the weak sense, and which approximate the unit in the sense of the quadratic metric.

---

[4] We may say: we replace the definition of the derivative as the limit of $h^{-1}[u(x+h) - u(x)]$ by a different limit process.

[5] Having proved the identity of both operators D^* in $\mathfrak{G}'$, we obtain the identity of both D in $\mathfrak{G}_\infty$ from the theorem of v. Neumann that a closed operator is the adjoint of its adjoint.

The theory of the operators D and D^* is presented in §§ 1-6 for any number n of the dimension and for any open set S.[6]

The §§ 7-14 are concerned with the application to the general symmetric elliptic differential operator of the second order, where the coefficients are only required to be continuous in the region S. We obtain the self-adjointness of this operator under two boundary conditions, which correspond to the vanishing of the function u or of its normal derivatives at the boundary of S.[7] In § 14 we apply the preceding theory to Schrödinger's operator.

In § 15 we deal with the second type of problems confining ourselves to the operators $D^*D = -\Delta$ and $-\Delta + q$ without boundary conditions. On the basis of the theory of §§ 1-6 we obtain the following result, which supplements well-known facts of potential theory:

Let n be the dimension of the x-space and introduce the number

$$m = \left[\frac{n}{2}\right] + 1.$$

Let $u(x)$ be a function on which the operations $\nabla_1 = D$, $\nabla_2 = D^*D$, $\nabla_3 = DD^*D, \cdots$ up to ∇_r can be applied such that $u, \nabla_1 u, \cdots, \nabla_r u$ are L^2-integrable. Then $u(x)$ is continuous and has continuous derivatives up to the order $r - m$ provided that $r \geqq m$.[8, 9, 10]

An immediate consequence is this: Let $u(x)$ be a characteristic function of the operator $D^*D = -\Delta$ under a certain boundary condition; (by char-

---

[6] No particular behaviour at the boundary will be required of the functions under consideration. For this, the concept of adjointness will be slightly modified.

It may be noted that a similar treatment is possible for any L^p-metric, where $p \geqq 1$, instead of Hilbert's L^2-metric.

[7] For a treatment of the manifold of all self-adjoint boundary conditions for the operator $-\frac{\partial}{\partial x} p \frac{\partial}{\partial x} - \frac{\partial}{\partial y} p \frac{\partial}{\partial y} + q$ and rectifiable boundary, see J. W. Calkin. *Proceedings of the National Academy of Sciences*, vol. 24 (January, 1938), 1, pp. 38-42 and a forthcoming note.

[8] An equivalent statement (except for $n = r = 1$) is this. Let u be a solution of $D^*Du = f$, where u and Du are L^2-integrable; then u possesses continuous derivatives up to the order r-m if f is L^2-integrable and admits the application of $\nabla_1, \nabla_2, \cdots$ in the above sense up to the order $r - 2$.

[9] It may be mentioned that a like proposition holds under any L^p-metric, $p \geqq 1$, with $m = \left[\frac{n}{p}\right] + 1$, except for the case that $p = 1$ and n is odd, where $m = n$.

The condition $r \geqq m$ is necessary for continuity since the discontinuous function $\mu = \log (c \mid \log \mid x \mid \mid)$ admits the above process $m - 1$ times.

[10] Cf. S. Sobolev. Sur quelques évaluations concernant les familles des fonctions ayant des dérivées à carré intégrable. *C. R. Acad. Sci. U R S S*, N. s. 1. 279-282 (1936). There a corresponding fact is stated where the applicability of all differentiations of the order r, instead of ∇_r, is required.

acteristic functions we mean not only eigenfunctions of point-eigenvalues λ but generally functions in the manifold of a projection belonging to any finite interval of the λ-axis). Then $u(x)$ has derivatives of every order. For, the above process can be carried on indefinitely, since u admits the application of D^*D and D^*Du is again a characteristic function; (in case u belongs to the point-eigenvalue λ this follows from $D^*Du = \lambda u$). In § 15 we shall establish the corresponding fact for the operator $D^*D + q(x)$, thus including Schrödinger's operator.

The present paper generalizes and supersedes the results of two previous papers [11] except those concerning point-spectra.

1. Definitions. Let $x = \{x_1, \cdots, x_n\}$ be the points of an n-dimensional space; by S we denote open regions of this space. If S' is a bounded region within S and if the boundary of S' is in S, too, we write $S' < S$.

Integrals over a domain S shall be written by $\int_S \cdots dx$ instead of $\int \cdots \int \cdots dx_1 \cdots dx_n$.

We denote by $\dot{\mathfrak{C}}_S$ the manifold of all continuous functions $u(x)$ defined in S to which a $S'' < S$ exists such that $u(x) = 0$ outside of S''; these functions shall be defined as identically zero outside of S.

We denote by $\mathfrak{L}_S$ the manifold of all functions $u(x)$ which are L^2-integrable in every region $S' < S$; we set

$$| u |^2_{S'} = \int_{S'} | u |^2 dx.$$

We denote by $\dot{\mathfrak{L}}_S$ the manifold of all $u(x)$ in $\mathfrak{L}_S$ which vanish outside of a region $S'' < S$; we set

$$| u |_S = | u |_{S''}.$$

Two functions $u(x)$, $v(x)$ in $\mathfrak{L}_S$ are called "equal" if $| u - v |_{S'} = 0$ for every $S' < S$.

For $u(x)$, $w(x)$ in $\mathfrak{L}_S$ the integral $\int_{S'} w(x)u(x)dx$ is defined and we have

$$\left| \int_{S'} w(x)u(x)dx \right| \leqq | w |_{S'} \, | u |_{S'}.$$

The same integral is defined for $w(x)$ in $\mathfrak{L}_S$, $u(x)$ in $\dot{\mathfrak{L}}_S$. If $|w|_{S''} = |w|_S$ then

$$\int_S w(x)u(x)dx = \int_{S''} w(x)u(x)dx.$$

---

[11] *Mathematische Annalen* Bd. 109 (1934), pp. 465-487, pp. 685-713. Bd. 110 (1935), p. 777.

If $u(x)$ is a function in $\mathfrak{L}_S$ such that

$$\int_S w(x)u(x)\,dx = 0$$

for all $w(x)$ in $\dot{\mathfrak{C}}_S$ then $u(x) = 0$.

2. The operator J_a. Let S' be a region $< S$. We denote by S'_a the region of all points y in S such that the region

$$|x_1 - y_1| < a, \cdots, |x_n - y_n| < a$$

contains at least a point x of S'. We assume a so small that

$$S'_{2a} < S.$$

We choose a function $e(t)$ which has derivatives of all orders, for which

$$e(t) \geqq 0 \text{ or } = 0 \text{ for } |t| \leqq 1 \text{ or } \geqq 1 \text{ respectively, and}$$

$$\int_{-\infty}^{+\infty} e(t)\,dt = 1.$$

We define the kernel

$$j_a(y-x) = a^{-n} e\left(\frac{y_1 - x_1}{a}\right) \cdots e\left(\frac{y_n - x_n}{a}\right)$$

and the integral operator J_a which takes the function $u(x)$ into the function

$$J_a u(x) = \int_{S'_a} j_a(y-x)u(y)\,dy.$$

If $u(x)$ is in $\mathfrak{L}_S$ then $J_a u(x)$ is a function in $\dot{\mathfrak{C}}_S$.

The operator J_a has the following properties:

(2.1) $$|J_a u|_S \leqq |u|_{S'_a}$$

for $u(x)$ in $\mathfrak{L}_S$.

(2.2)[12] $$|J_a u - u|_{S'} \to 0, \text{ as } a \to 0$$

for $u(x)$ in $\mathfrak{L}_S$.

(2.3) $$|J_a u - u|_S \to 0, \text{ as } a \to 0$$

for $u(x)$ in $\dot{\mathfrak{L}}_S$, if $S''_a < S'$ where $|u|_S = |u|_{S''}$.

For $u(x)$ in $\mathfrak{L}_S$ we have the relation

(2.4) $$|J_a u(x)|^2 \leqq \int_{S'_a} j_a(y-x)\,|u(y)|^2\,dy;$$

[12] For the case $n = 1$ this is contained in the results of K. Ogura, *Tôhôku Mathematical Journal*, vol. 16 (1919), pp. 118-125 (Theorem II).

for we have

$$| J_a u(x) | = | \int_{S'_a} | j_a(y-x) |^{\frac{1}{2}} | j_a(y-x) |^{\frac{1}{2}} u(y) dy |$$
$$\leqq | \int_{S'_a} j_a(y-x) dy |^{\frac{1}{2}} | \int_{S'_a} j_a(y-x) | u(y) |^2 dy |^{\frac{1}{2}}$$
$$\leqq | \int_{S'_a} j_a(y-x) | u(y) |^2 dy |^{\frac{1}{2}}.$$

From (2.4) we deduce

$$\int_S | J_a u(x) |^2 dx \leqq \int_S \int_{S'_a} j_a(y-x) | u(y) |^2 dy dx$$
$$= \int_{S'_a} \int_S j_a(y-x) dx \, | u(y) |^2 dy = \int_{S'_a} | u(y) |^2 dy,$$

and thus the inequality (2.1).

To prove the relation (2.2) we observe that $u(x)$ can be approximated by functions $u_b(x)$ in $\dot{\mathfrak{C}}_S$ such that $| u - u_b |_{S'_a} \to 0$ as $b \to 0$; in view of inequality (2.1) we get $| J_a(u - u_b) |_S \to 0$, as $b \to 0$. Therefore it is sufficient to prove $| J_a u_b - u_b |_S \to 0$ as $a \to 0$. For x in S', we get

$$| J_a u_b(x) - u_b(x) | = | \int_{S'_a} j_a(y-x) [u_b(y) - u_b(x)] dy |$$
$$\leqq \text{l. u. b.} \, | u_b(y) - u_b(x) | \quad \text{for} \quad | x_1 - y_1 | < a, \cdots, | x_n - y_n | < a.$$

Since $u_b(x)$ is continuous the right hand term tends to zero as $a \to 0$; therefore

$$| J_a u_b - u_b |_{S'} = \int_{S'} | J_a u_b(x) - u_b(x) |^2 dx \to 0.$$

If $u(x)$ is in $\dot{\mathfrak{L}}_S$, $| u |_S = | u |_{S''}$, $S''_a < S'$, we have $J_a u(x) = 0$ outside of S'; therefore $| J_a u - u |_S = | J_a u - u |_{S'}$. Thus we obtain (2.3).

3. Systems. By $n + 1$ functions $v_0(x), v_1(x), \cdots, v_n(x)$ we form the system

$$v'(x) = \{ v_0(x), v_1(x), \cdots, v_n(x) \}.$$

The manifold of all such systems of functions in $\dot{\mathfrak{C}}$, $\mathfrak{L}$, $\dot{\mathfrak{L}}$ shall be denoted by $\dot{\mathfrak{C}}'$, $\mathfrak{L}'$, $\dot{\mathfrak{L}}'$ respectively. For $v'(x)$ in $\mathfrak{L}'_S$ we set

$$| v' |_S^2 = | v_0 |_S^2 + \cdots + | v_n |_S^2.$$

For $w'(x)$, $v'(x)$ in $\mathfrak{L}'_S$ and for $w'(x)$ in $\dot{\mathfrak{C}}'_S$, $v'(x)$ in $\mathfrak{L}'_S$ we set

$$\int_{S'} w'(x) v'(x) dx = \int_{S'} w_0(x) v_0(x) dx + \cdots + \int_{S'} w_n(x) v_n(x) dx.$$

Finally we set

$$J_a v'(x) = J_a\{v_0(x), \cdots, v_n(x)\} = \{J_a v_0(x), \cdots, J_a v_n(x)\}.$$

4. Differential operators in the interior. Let $\dot{\mathfrak{D}}_S$ and $\dot{\mathfrak{D}}'_S$ be the spaces of all functions $u(x)$ in $\dot{\mathfrak{C}}_S$ and $v'(x)$ in $\dot{\mathfrak{C}}'_S$ respectively which have continuous derivatives

$$D_\mu u(x) = \frac{\partial}{\partial x_\mu} u(x); \quad D_\mu v_\nu(x) = \frac{\partial}{\partial x_\mu} v_\nu(x); \qquad (\mu, \nu = 1, \cdots, n).$$

By D we denote the operator which takes $u(x)$ in $\dot{\mathfrak{D}}_S$ into the function in $\dot{\mathfrak{C}}'_S$

$$Du(x) = \{\delta u(x), D_1 u(x), \cdots, D_n u(x)\},$$

where δ is an arbitrary fixed number (which may be zero). By D^* we denote the operator which takes the function $v'(x)$ in $\dot{\mathfrak{D}}'_S$ into the function in $\dot{\mathfrak{C}}_S$

$$D^* v'(x) = \delta v_0(x) - D_1 v_1(x) - \cdots - D_n v_n(x).$$

For $u(x)$ in $\dot{\mathfrak{D}}_S$ and $v'(x)$ in $\dot{\mathfrak{D}}'_S$ we have

$$\int_S v'(x) Du(x)\, dx = \int_S D^* v'(x) \cdot u(x)\, dx. \tag{4.0}$$

5. Operator D in $\mathfrak{K}_S$. We now extend the operator D from $\dot{\mathfrak{D}}_S$ to a space $\mathfrak{K}_S \subset \mathfrak{L}_S$. We define $\mathfrak{K}_S$ as the manifold of all functions in $\mathfrak{L}_S$ for which a function $u'(x)$ in $\mathfrak{L}_S$ exists such that

$$\int_S w'(x) u'(x)\, dx = \int_S D^* w'(x) \cdot u(x)\, dx$$

for all $w'(x)$ in $\dot{\mathfrak{D}}'_S$. If $u(x) = 0$ then $u'(x) = 0$, also. For, since $u(x) = 0$ in S we have $\int_S w' u' dx = 0$ for all w' in $\dot{\mathfrak{D}}'_S$; from this one obtains [13] $\int_{S'} |u'|^2 dx = 0$, hence $u' = 0$ in each $S' < S$. In $\mathfrak{K}_S$, therefore, we can define a linear operator D by

$$Du(x) = \{\delta u(x), D_1 u(x), \cdots, D_n u(x)\} = u'(x),$$

which is evidently an extension of the operator D in $\dot{\mathfrak{D}}_S$.

We shall speak of a differential operator defined in this way as being an

[13] *E. g.* Set $w' = J_a u'$ and use (2.2) and (2.1).

extension in the "weak" sense. Immediate consequences of this definition are the properties:[14]

5.1. If $u(x)$ in $\mathfrak{R}_S$, $w'(x)$ in $\dot{\mathfrak{D}}'_S$, then

$$\int_S w'(x) Du(x)\, dx = \int_S D^* w'(x) \cdot u(x)\, dx. \tag{5.1}$$

5.2. If $u(x)$ in $\mathfrak{L}_S$, $u'(x)$ in $\mathfrak{L}'_S$ is such that

$$\int_S w'(x) u'(x)\, dx = \int_S D^* w'(x) \cdot u(x)\, dx \tag{5.2}$$

for all $w'(x)$ in $\dot{\mathfrak{D}}'_S$ then $u(x)$ is in $\mathfrak{R}_S$ and $Du(x) = u'(x)$.

As a matter of fact the operator D in $\mathfrak{R}_S$ can be shown to be an extension in a "strong" sense also, for we have

5.3. If $u(x)$ in $\mathfrak{L}_S$, $u'(x)$ in $\mathfrak{L}'_S$ has the property that for every $S' < S$ a sequence $u_a(x)$ in $\dot{\mathfrak{D}}_S$ exists such that

$$| u_a - u |_{S'} + | Du_a - u' |_{S'} \to 0 \tag{5.3}$$

then $u(x)$ is in $\mathfrak{R}_S$ and $Du(x) = u'(x)$.

5.4. For every function $u(x)$ in $\mathfrak{R}_S$ and region $S' < S$ there exists a sequence $u_a(x)$ in $\dot{\mathfrak{D}}_S$ such that

$$| u_a - u |_{S'} + | Du_a - Du |_{S'} \to 0 \text{ as } a \to 0.^{15,16} \tag{5.4}$$

Property 5.3 is an immediate consequence of 5.2: for every $w'(x)$ in $\dot{\mathfrak{D}}'_S$ there is a $S' < S$ such that $w'(x) = 0$ outside of S'; now we choose a sequence $u_a(x)$ in $\dot{\mathfrak{D}}_S$ such that (5.3) holds, from (4.0), for $u = u_a$, $v' = w'$, we obtain (5.2) as $a \to 0$; thus, according to 5.3, $u(x)$ is in $\mathfrak{R}_S$ and $Du(x) = u'(x)$.

In order to prove 5.4 we take the operator J_a. We observe that $J_a u(x)$ is in $\dot{\mathfrak{D}}_S$ if $u(x)$ is in $\mathfrak{L}_S$. We then state the basic

---

[14] They do not exactly express that D in $\mathfrak{R}_S$ is adjoint to D^* in $\dot{\mathfrak{D}}_S$ since the functions $u(x)$ in $\mathfrak{R}_S$ are not restricted to be in the Hilbert space defined by the unit form $\int_S | u |^2 dx$.

[15] These properties do not exactly express that D in $\mathfrak{R}_S$ is the closure of D in $\dot{\mathfrak{D}}_S$ for the same reason as in footnote [14].

[16] 5.3 states that the strong extension is contained in the weak extension, while 5.4 states the converse.

LEMMA 5.1. *If $u(x)$ is in $\mathfrak{K}_S$ then*

$$DJ_a u(x) = J_a Du(x)$$

for x in S'.

To prove it we observe that, for $\mu \neq 0$,

$$D_\mu J_a u(x) = \frac{\partial}{\partial x_\mu} \int_{S'_a} j_a(y-x)u(y)dy$$
$$= -\int_{S'_a} \frac{\partial}{\partial y_\mu} j_a(y-x)u(y)dy.$$

Since $j_a(y-x)$, when considered as a function of y, is in $\dot{\mathfrak{D}}_S$ if x is in S' we can define a function $w'(y)$ in $\dot{\mathfrak{D}}'_S$ by $w_\mu(y) = j_a(y-x)$, $w_\nu(y) = 0$ if $\nu = \mu$. Then we get

$$DJ_a u(x) = \int_{S'_a} D^* w'(y) \cdot u(y)dy.$$

In view of 5.1 this is

$$= \int_{S'_a} w'(y)Du(y)dy = \int_{S'_a} j_a(y-x)D_\mu u(y)dy = J_a D_\mu u(x)$$

and thus $DJ_a u(x) = J_a Du(x)$.

From this lemma we deduce 5.4 immediately. For if $u(x)$ in $\mathfrak{K}_S$, we obtain from (2.2)

$$|J_a u - u|_{S'} \to 0$$

and [17]

$$|DJ_a u - Du|_{S'} = |J_a Du - Du|_{S'} \to 0 \text{ as } a \to 0.$$

Finally we define the space $\dot{\mathfrak{K}}_S$ of all $u(x)$ in $\mathfrak{K}_S$ which are in $\dot{\mathfrak{L}}_S$ and for which $Du(x)$ is in $\dot{\mathfrak{L}}'_S$. Then we have 5.5 for every function $u(x)$ in $\dot{\mathfrak{K}}_S$ a sequence u_a in $\dot{\mathfrak{D}}_S$ exists such that

$$(5.5) \qquad |u_a - u|_S + |Du_a - Du|_S \to 0, \text{ as } a \to 0.$$

We obtain this property from Lemma 5.1, on using (2.3) and choosing S'' such that

$$|u|_S = |u|_{S''}, \quad |Du|_S = |Du|_{S''}, \quad S' > S''.$$

6. Operator D^* in $\mathfrak{K}_S$. In a corresponding way we extend the operator D^* from $\dot{\mathfrak{D}}'_S$ to a space $\mathfrak{K}_S$. We define $\mathfrak{K}_S$ as the manifold of all $v(x)$ in $\mathfrak{L}'_S$ for which a $v(x)$ in $\mathfrak{L}_S$ exists such that

[17] Cf. for $n = 1$ K. Ogura (*l. c.* 12), Theorem VII.

$$\int_S w(x)v(x)dx = \int_S Dw(x) \cdot v'(x)dx$$

for all $w(x)$ in $\dot{\mathfrak{D}}_S$. As in § 5 we conclude that $v'(x) = 0$ implies $v(x) = 0$. Thus we can define a linear operator D^* in $\mathfrak{K}'_S$ by

$$D^*v'(x) = v(x),$$

which is obviously an extension of D^* in D'_S. We call it an extension in the "weak" sense. Immediate consequences of this definition are the properties[18]

6.1. If $v'(x)$ in $\mathfrak{K}'_S$, $w(x)$ in $\dot{\mathfrak{D}}_S$, then

$$\int_S w(x)D^*v'(x)dx = \int_S Dw(x) \cdot v'(x)dx. \tag{6.1}$$

6.2. If $v'(x)$ in $\mathfrak{L}'_S$, $v(x)$ in $\mathfrak{L}_S$ is such that

$$\int_S w(x)v(x)dx = \int_S Dw(x)v'(x)dx \tag{6.2}$$

for all $w(x)$ in $\dot{\mathfrak{D}}_S$ then $v'(x)$ is in $\mathfrak{K}'_S$ and $D^*v'(x) = v(x)$.

Further the operator D^* in $\mathfrak{K}'_S$ has the following properties[19, 20]

6.3. If the functions $v'(x)$ in $\mathfrak{L}'_S$, $v(x)$ in $\mathfrak{L}'_S$ have the property that for every region $S' < S$ a sequence $v'_a(x)$ in $\dot{\mathfrak{D}}'_S$ exists such that

$$| v'_a - v' |_{S'} + | D^*v'_a - v |_{S'} \to 0, \text{ as } a \to 0, \tag{6.3}$$

then $v'(x)$ is in $\mathfrak{K}'_S$ and $D^*v'(x) = v(x)$.

6.4. For every $v'(x)$ in $\mathfrak{K}'_S$ and region $S' < S$ a sequence $v'_a(x)$ in $\dot{\mathfrak{D}}'_S$ exists such that

$$| v'_a - v' |_{S'} + | D^*v'_a - D^*v' |_{S'} \to 0, \text{ as } a \to 0. \tag{6.4}$$

Property 6.3 is an immediate consequence of 5.2. Property 6.4 is obtained, on setting $v'_a(x) = J_a v'(x)$, from

Lemma 6.1. *If $v'(x)$ is in $\mathfrak{K}'_S$ then*

$$D^*J_a v'(x) = J_a D^*v'(x).$$

This lemma is proved in the same way as Lemma 5.1 on applying (6.1) to $w(y) = j_a(y - x)$.

Finally we define $\dot{\mathfrak{K}}'_S$ to be the space of all $v'(x)$ in $\mathfrak{K}'_S$ which are in $\dot{\mathfrak{L}}'_S$ and for which $D^*v'(x)$ is in $\dot{\mathfrak{L}}_S$. Then we have

[18, 19, 20] Remarks corresponding to [14, 15, 16] may be applied respectively.

6.5. To every function in $v'(x)$ in $\dot{\mathfrak{K}}'_S$ a sequence $v'_a(x)$ in $\dot{\mathfrak{D}}'_S$ exists such that

$$(6.5) \qquad | v'_a - v' |_S + | D^*v'_a - D^*v' |_S \to 0, \text{ as } a \to 0.$$

7. Closed spaces $\mathfrak{H}$ and $\mathfrak{H}'$. In what follows we omit the subscript S.

Let $r(x)$ be a positive continuous function defined in S. Let R and R^{-1} be the operators which take a function $u(x)$ in $\mathfrak{L}$ respectively into the functions in $\mathfrak{L}$

$$Ru(x) = r(x)u(x), \quad R^{-1}u(x) = r(x)^{-1}u(x).$$

For $u(x)$, $w(x)$ in $\mathfrak{L}$, $S' < S$, the bilinear form

$$\int_{S'} w(x)Ru(x)dx$$

is defined and we have

$$\int_{S'} u(x)Ru(x)dx > 0, \text{ for } u(x) \neq 0 \text{ in } S'.$$

We define

$$(u\boldsymbol{R}u) = \int_S u(x)Ru(x)dx = \text{l. u. b.}_{S'<S} \int_{S'} u(x)Ru(x)dx$$

and introduce the manifold $\mathfrak{H}$ of all $u(x)$ in $\mathfrak{L}$ for which $(u\boldsymbol{R}u) < \infty$. For $u(x)$, $w(x)$ in $\mathfrak{H}$ the bilinear form

$$(w\boldsymbol{R}u) = \int_S w(x)Ru(x)dx$$

is defined. $\mathfrak{H}$ is a Hilbert space and contains $\dot{\mathfrak{C}}$ densely.

Let $p_{\mu\nu}(x)$; $(\mu, \nu = 0, \cdots, n)$ be continuous functions defined in S such that at every x in S the matrix $p_{\mu\nu}$ is symmetric and positive definite. Let P be the operator which takes every function $v'(x) = \{v_0(x), \cdots, v_n(x)\}$ in $\mathfrak{L}'$ into the function in $\mathfrak{L}'$

$$Pv'(x) = \{ \sum_{\nu=0}^{n} p_{0\nu}(x)v_\nu(x), \cdots, \sum_{\nu=0}^{n} p_{n\nu}(x)v_\nu(x)\}.$$

For $v'(x)$, $w'(x)$ in $\mathfrak{L}'$, $S' < S$, the bilinear form

$$\int_{S'} w'(x)Pv'(x)dx$$

is defined and we have

$$\int_{S'} v'(x)Pv'(x)dx > 0, \text{ as } v'(x) \neq 0 \text{ in } S'.$$

We define

$$(v'\boldsymbol{P}v') = \int_S v'(x)Pv'(x)\,dx = \text{l. u. b.}_{S'<S} \int_{S'} v'(x)Pv'(x)\,dx$$

and introduce the manifold $\mathfrak{H}'$ of all $v'(x)$ in $\mathfrak{L}$ for which $(v'\boldsymbol{P}v') < \infty$. For $v'(x)$, $w'(x)$ in $\mathfrak{H}'$ the bilinear form

$$(w'\boldsymbol{P}u') = \int_S w'(x)Pv'(x)\,dx$$

is defined. $\mathfrak{H}'$ is a Hilbert space and contains $\dot{\mathfrak{C}}'$ densely.

We denote by $P^{-1}\mathfrak{K}'$, $P^{-1}\dot{\mathfrak{K}}'$, $P^{-1}\dot{\mathfrak{D}}'$ the space of all functions $v'(x)$ for which $Pv'(x)$ is in $\mathfrak{K}'$, $\dot{\mathfrak{K}}'$, $\dot{\mathfrak{D}}'$ respectively. In all these spaces the operator

$$E = R^{-1}D^*P$$

is defined; it produces functions in $\mathfrak{L}$, $\dot{\mathfrak{L}}$, $\dot{\mathfrak{C}}$ respectively.

8. The spaces $\mathfrak{G}_0$ and $\mathfrak{G}_\infty$. We denote by $\mathfrak{G}_0$ the space of all $u(x)$ in $\mathfrak{H}$ which are in $\mathfrak{K}$ and for which $Du(x)$ is in $\mathfrak{H}'$. In this space the unit form

$$\| u \|^2 = (u\boldsymbol{R}u) + (Du\boldsymbol{P}Du)$$

is defined. We have

THEOREM 8.1. *The space $\mathfrak{G}_0$ is complete with respect to $\| u \|$.*

This can be deduced from 5.3 by well-known reasoning.

Obviously $\mathfrak{G}_0$ contains $\dot{\mathfrak{K}}$ (cf. § 5). We denote by $\mathfrak{G}_\infty$ the closure of $\dot{\mathfrak{K}}$ with respect to $\| u \|$.[21]

We have

THEOREM 8.2. *The space $\mathfrak{G}_\infty$ contains $\dot{\mathfrak{D}}$ densely.*[22]

It suffices to prove that to every function $u(x)$ in $\dot{\mathfrak{K}}$ a sequence $u_a(x)$ in $\dot{\mathfrak{D}}$ exists such that $\| u_a - u \| \to 0$ as $a \to 0$.

We choose a $S^0 < S$ such that u is in $\dot{\mathfrak{K}}_{S^0}$ and apply 5.5 to S^0 instead of S; since the coefficients of R and P are bounded in S^0 and u_a in $\dot{\mathfrak{D}}_{S^0}$ is in $\dot{\mathfrak{D}}_S$, too, we obtain the statement.

---

[21] The condition for a function of $\mathfrak{G}_0$ to be in $\mathfrak{G}_\infty$ is a boundary condition; for if u_1 is in $\mathfrak{G}_\infty$ and u_2 in $\mathfrak{G}_0$ differs from u_1 only in the interior then u_2 is in $\mathfrak{G}_\infty$, too, because $u_2 - u_1$ is in $\dot{\mathfrak{K}}$ (cf. § 5, p. 8) and therefore in $\mathfrak{G}_\infty$.

[22] It would be possible to define $\mathfrak{G}_\infty$ as the closure of $\dot{\mathfrak{D}}$. Then, instead of Th. 8.2, we would prove that $\dot{\mathfrak{K}}$ is in $\mathfrak{G}_\infty$ and thus the condition defining $\mathfrak{G}_\infty$ is a boundary condition.

9. The spaces $\mathfrak{G}'_0$ and $\mathfrak{G}'_\infty$. We denote by $\mathfrak{G}'_0$ the space of all $v'(x)$ in $\mathfrak{H}'$ which are in $P^{-1}\mathfrak{K}'$ and for which $Ev' = R^{-1}D^*Pv'$ is in $\mathfrak{H}$. In this space the unit form

$$\| v' \|^2 = (v'\boldsymbol{P}v') + (Ev'\boldsymbol{R}Ev')$$

is defined. As a consequence of 6. 3 we have

THEOREM 9. 1. *The space $\mathfrak{G}'_0$ is complete with respect to $\| v' \|$.*

Obviously $\mathfrak{G}'_0$ contains the space $P^{-1}\dot{\mathfrak{K}}'$ of all functions $v'(x)$ for which $Pv'(x)$ is in $\dot{\mathfrak{K}}'$. We introduce the closure $\mathfrak{G}'_\infty$ of $P^{-1}\dot{\mathfrak{K}}'$ with respect to $\| v' \|$.[23] We state

THEOREM 9. 2. *The space $\mathfrak{G}'_\infty$ contains $P^{-1}\dot{\mathfrak{D}}'$ densely.*[24]

This follows immediately from 6. 5, as § 2 had followed from 5. 2.

10. The operators D in $\mathfrak{G}_0$ and E in $\mathfrak{G}'_\infty$. The operator D defined in $\mathfrak{G}_0$ produces functions in $\mathfrak{H}'$ while $E = R^{-1}D^*P$ is defined in $\mathfrak{G}'_\infty$ and produces functions in $\mathfrak{H}$. Between these operators the following relations hold.

10. 1. If $v'(x)$ is in $\mathfrak{G}'_\infty$, $u(x)$ in $\mathfrak{G}_0$, then

$$(uREv') = (Du\boldsymbol{P}v'). \tag{10. 1}$$

10. 2. If $u(x)$ in $\mathfrak{H}$, $u'(x)$ in $\mathfrak{H}'$ are such that

$$(u\boldsymbol{R}Ew') = (u'\boldsymbol{P}w') \tag{10. 2}$$

for all $w'(x)$ in $P^{-1}\dot{\mathfrak{D}}'$ then $u(x)$ is in $\mathfrak{G}_0$ and $Du(x) = u'(x)$.
10. 3 If $v'(x)$ in $\mathfrak{H}'$, $v(x)$ in $\mathfrak{H}$ are such that

$$(w\boldsymbol{R}v) = (Dw\boldsymbol{P}v') \tag{10. 3}$$

for all $w(x)$ in $\mathfrak{G}_0$ then $v'(x)$ is in $\mathfrak{G}'_\infty$ and $v(x) = Ev'(x)$.

10. 1 and 10. 2 imply that D in $\mathfrak{G}_0$ is adjoint to E in $\mathfrak{G}'_\infty$. 10. 1 and 10. 3 state that E in $\mathfrak{G}'_\infty$ is adjoint to D in $\mathfrak{G}'_0$.

10. 1 follows from 5. 1 and Th. 9. 2; 10. 2 follows from 5. 2. 10. 3 is a consequence of a basic theorem of von Neumann [25] which states that a closed operator is the adjoint of its adjoint. It seems desirable to repeat his reasoning for our case.

---

[23] Cf. [21]. [24] Cf. [22]. [25] Cf. [2].

Let $[\mathfrak{H}, \mathfrak{H}']$ be the space of all pairs $[u(x), v'(x)]$ of $u(x)$ in $\mathfrak{H}$, $v'(x)$ in $\mathfrak{H}'$; it is a Hilbert space with respect to the unit form

$$(u\boldsymbol{R}u) + (v'\boldsymbol{P}v').$$

Let $[E\mathfrak{G}'_\infty, \mathfrak{G}'_\infty]$ be the subspace of all pairs $[Ev'(x), v'(x)]$ where $v'(x)$ is in $\mathfrak{G}'_\infty$. This space is closed according to the definition of $\mathfrak{G}'_\infty$. Let $[\mathfrak{G}_0, -D\mathfrak{G}_0]$ be the subspace of all pairs $[u(x), -Du(x)]$ where $u(x)$ is in $\mathfrak{G}_0$. According to 10. 1 and 10. 2 this space consists of all elements of $[\mathfrak{H}, \mathfrak{H}']$ which are orthogonal to $[E\mathfrak{G}'_\infty, \mathfrak{G}'_\infty]$. Since $[\mathfrak{H}, \mathfrak{H}']$ is a Hilbert space the closed space $[E\mathfrak{G}'_\infty, \mathfrak{G}'_\infty]$ contains all elements of $[\mathfrak{H}, \mathfrak{H}']$ which are orthogonal to $[\mathfrak{G}_0, -D\mathfrak{G}_0]$. That is exactly the statement of 10. 3.

11. The operator D in $\mathfrak{G}_\infty$ and E in $\mathfrak{G}'_0$. Between the operators D in $\mathfrak{G}_\infty$ and E in $\mathfrak{G}'_0$ the following relations hold

11. 1. If $u(x)$ is in $\mathfrak{G}_\infty$, $v'(x)$ in $\mathfrak{G}'_0$ then

$$(v'\boldsymbol{P}Du) = (Ev'\boldsymbol{R}u). \tag{11. 1}$$

11. 2. If $v'(x)$ in $\mathfrak{H}'$ and $v(x)$ in $\mathfrak{H}$ are such that

$$(v'\boldsymbol{P}Dw) = (v\boldsymbol{R}w) \tag{11. 2}$$

for all $w(x)$ in $\dot{\mathfrak{D}}$ then $v'(x)$ is in $\mathfrak{G}'_0$ and $Ev'(x) = v(x)$.

11. 3. If $u(x)$ in $\mathfrak{H}$ and $u'(x)$ in $\mathfrak{H}'$ are such that

$$(w'\boldsymbol{P}u') = (Ew'\boldsymbol{R}u) \tag{11. 3}$$

for all $w'(x)$ in $\mathfrak{G}'_0$ then $u(x)$ is in $\mathfrak{G}_\infty$ and $u'(x) = Du(x)$.

11. 1 and 11. 2 includes the fact that E in $\mathfrak{G}'_0$ is adjoint to D in $\mathfrak{G}_\infty$. They follow from 6. 1 and Th. 8. 2. 11. 1 and 11. 3 state that D in $\mathfrak{G}_\infty$ is adjoint to E in $\mathfrak{G}'_0$; this is again a consequence of the theorem of von Neumann.

12. Self adjoint operator ED in $\mathfrak{F}$. Let $\mathfrak{F}_\infty$ be the space of all functions $u(x)$ in $\mathfrak{G}_\infty$ for which $Du(x)$ is in $\mathfrak{G}'$. Let $\mathfrak{F}_0$ be the space of all functions $u(x)$ in $\mathfrak{G}_0$ for which $Du(x)$ is in $\mathfrak{G}'_\infty$. The operator ED is defined in $\mathfrak{F}_\infty$ and in $\mathfrak{F}_0$ and produces functions in $\mathfrak{H}$.

It is convenient to denote by $\mathfrak{G}$, $\mathfrak{G}'$, $\mathfrak{F}$ either $\mathfrak{G}_\infty$, $\mathfrak{G}'_0$, $\mathfrak{F}_\infty$ or $\mathfrak{G}_0$, $\mathfrak{G}'_\infty$, $\mathfrak{F}_0$ respectively.

It is by no means evident, that the spaces $\mathfrak{F}$ are dense in $\mathfrak{H}$, and that they contain functions other than $u(x) = 0$. But according to a theorem of

von Neumann (l. c.[2]), $\mathfrak{F}$ is dense and ED in $\mathfrak{F}$ is self-adjoint (hypermaximal). Explicitly we have

$\mathfrak{F}$ is dense in $\mathfrak{H}$ with respect to $(u\boldsymbol{R}u)$.

12.1. If $u(x)$ and $w(x)$ are in $\mathfrak{F}$ then

$$(w\boldsymbol{R}EDu) = (EDw\boldsymbol{R}u). \tag{12.1}$$

12.2. If $u(x)$, $v(x)$ are functions in $\mathfrak{H}$ such that

$$(w\boldsymbol{R}v) = (EDw\boldsymbol{R}u) \tag{12.2}$$

for all $w(x)$ in $\mathfrak{F}$ then $u(x)$ is in $\mathfrak{F}$ and $EDu(x) = v(x)$.

Further we have

$\mathfrak{F}$ is dense in $\mathfrak{G}$ with respect to $(Du\boldsymbol{P}Du)$.

12.3. If $u(x)$ is in $\mathfrak{F}$, $w(x)$ in $\mathfrak{G}$, then

$$(w\boldsymbol{R}EDu) = (Dw\boldsymbol{P}Du). \tag{12.3}$$

12.4. If $u(x)$ in $\mathfrak{G}$, $v(x)$ in $\mathfrak{H}$ are such that

$$(w\boldsymbol{R}v) = (Dw\boldsymbol{P}Du) \tag{12.4}$$

for all $w(x)$ in $\mathfrak{G}$ then $u(x)$ is in $\mathfrak{F}$ and $EDu(x) = v(x)$.

We prove the preceding statement in a way,[26] which differs slightly from von Neumann's reasoning. From 10.1 or 11.1 we obtain 12.3; from this and 10.1 or 11.1 we obtain 12.1. We deduce from 12.2 that $\mathfrak{F}$ is dense in $\mathfrak{H}$; for if $wRv = 0$ for all w in $\mathfrak{F}$ then (12.2) is satisfied by $u(x) = 0$; hence $v = EDu = 0$. Further, we deduce from 12.2 that $\mathfrak{F}$ is dense in $\mathfrak{G}$; for if $(Dw\boldsymbol{P}Du) = 0$ for all w in $\mathfrak{F}$ and some u in $\mathfrak{G}$ we find from 12.3 that (12.2) is satisfied by $v = 0$; hence u is in $\mathfrak{F}$; therefore $(Du\boldsymbol{P}Du) = 0$ and $u = 0$.

We thus need prove only 12.2. We first state

LEMMA 12.1. *To every function $h(x)$ in $\mathfrak{H}$ there is a function $f(x)$ in $\mathfrak{F}$ such that $(ED + 1)f = h$.*

To prove this lemma we observe that $(w\boldsymbol{R}h)$ is a linear form for functions $w(x)$ in $\mathfrak{G}$ which is bounded with respect to the unit form $(Dw\boldsymbol{P}Dw) + (w\boldsymbol{R}w)$. Therefore a function f in $\mathfrak{G}$ exists such that

$$(Dw\boldsymbol{P}Df) + (w\boldsymbol{R}f) = (w\boldsymbol{R}h)$$

---

[26] Cf. the corresponding reasoning in K. Friedrichs, *Mathematische Annalen* Bd. 109, S. 465, 1934, H. Freudenthal, *Proc. Kon. Acad. Wet Amsterdam*, 39, N. 7, 1936.

for all w in $\mathfrak{G}$. From 10. 3 or 11. 2 we obtain that f is in $\mathfrak{F}$ and $(ED+1)f=h$.

Before applying this lemma we remark that 12. 2 is evidently equivalent to 12. 2′. If $u(x)$, $v_1(x)$ are functions in $\mathfrak{H}$ such that

$$(w\boldsymbol{R}v_1) = ((ED+1)w\boldsymbol{R}u) \tag{12.2'}$$

for all w in $\mathfrak{F}$ then $u(x)$ is in $\mathfrak{F}$ and $(ED+1)u(x) = v_1(x)$.

Let $u(x)$, $v_1(x)$ be such functions. In view of Lemma 12. 1 there is a function $u_1(x)$ in $\mathfrak{F}$ such that $(ED+1)u_1(x) = v_1(x)$. Hence, from (12. 2′), $((ED+1)w\boldsymbol{R}u) = (w\boldsymbol{R}(ED+1)u_1) = ((ED+1)wRu_1)$ in view of 12. 1. Thus $((ED+1)w\boldsymbol{R}(u_1-u))=0$. On applying the lemma again we may choose the function $w(x)$ such that $(ED+1)w(x) = u_1(x) - u(x)$; thus we get

$$((u_1-u)\boldsymbol{R}(u_1-u)) = 0, \text{ hence } u = u_1 \text{ in } \mathfrak{F}.$$

13. Modification of the matrix *P*. Let $q_1(x), \cdots, q_n(x)$ be continuous functions in S which have continuous first derivatives. Then we define a "difference operator" ΔP by the matrix

$$\begin{array}{lll}
\Delta p_{\mu\nu}(x) = 0 & & \mu \neq 0, \quad \nu \neq 0 \\
\Delta p_{0\nu}(x) = q_\nu(x) & & \nu \neq 0 \\
\Delta p_{\mu 0}(x) = q_\mu(x) & & \mu \neq 0 \\
\Delta p_{00}(x) = (D_1 q_1(x) + \cdots + D_n q_n(x)). & &
\end{array}$$

On setting

$$\Delta E = R^{-1}D^*\Delta P$$

we have

LEMMA 13. 1. *If ΔP is a difference operator, $u(x)$ in $\mathfrak{R}$ then $\Delta PDu(x)$ is in $\mathfrak{R}'$ and $\Delta EDu(x) = 0$.*

We set $q' = (0, q_1, \cdots, q_n)$; then we have

$$\Delta PDu(x) = \{-D^*q' \cdot u + q'Du, q_1 u, \cdots, q_n u\}.$$

Let $w(x)$ be a function in $\dot{\mathfrak{D}}$ then we obtain from 5. 1

$$\int_S w[D^*q' \cdot u - q'Du]dx = \int_S [wD^*q' - D^*wq']udx = \int_S Dw \cdot q'udx.$$

Hence 6. 2 shows that $q'u$ is in $\mathfrak{R}'$ and $D^*q'u = D^*q' \cdot u - q'Du$; therefore $\Delta PDu(x)$ is in $\mathfrak{R}'$, too, and $D^*\Delta PDu(x) = 0$. By the matrix

$$r_{00}(x) = r(x), \qquad r_{\mu\nu}(x) = 0 \text{ otherwise},$$

we introduce an operator which we denote also by R.

Then we term the difference operator ΔP "admissible" to P if a constant c exists such that the operator

$$P + \Delta P + cR$$

is positive definite in every point of S. For this operator we define the spaces $\mathfrak{G}(\Delta P)$, $\mathfrak{G}_\infty(\Delta P)$, $\mathfrak{F}(\Delta P)$, $\mathfrak{F}_\infty(\Delta P)$[27] in the same way as the spaces $\mathfrak{G}_0$, $\mathfrak{G}_\infty$, $\mathfrak{F}_0$, $\mathfrak{F}_\infty$ for the operator P.

If ΔP vanishes in a neighborhood of the boundary it is admissible and the spaces $\mathfrak{G}_0$, $\mathfrak{G}_\infty$, $\mathfrak{F}_0$, $\mathfrak{F}_\infty$ are not changed. Thus the modification of these spaces depends on the behavior of ΔP only at the boundary of S. Therefore we can say the operators ED in $\mathfrak{F}(\Delta P)$ differ only in their behavior at the boundary and the manifold of admissible ΔP represents a manifold of boundary conditions. But this does not concern the operator ED in $\mathfrak{F}_\infty$; for we have

THEOREM 13. 1. $\mathfrak{F}_\infty(\Delta P) = \mathfrak{F}_\infty$.

According to Lemma 13. 1 we then have

$$(E + \Delta E)Du(x) = EDu(x).$$

We first prove $\mathfrak{F}_\infty(\tfrac{1}{2}\Delta P) = \mathfrak{F}_\infty$. From

$$(13.1) \qquad (v', Pv') + (v', [P + \Delta P + cR]v') = 2(v', [P + \tfrac{1}{2}\Delta P + \tfrac{1}{2}cR]v')$$

we see that the difference operator $\frac{1}{2}\Delta P$ is admissible and that $\mathfrak{G}_\infty \supset \mathfrak{G}_\infty(\frac{1}{2}\Delta P)$. In view of Lemma 13. 1 we then obtain

$$\mathfrak{F}_\infty \supset \mathfrak{F}_\infty(\tfrac{1}{2}\Delta P).$$

But since ED in $\mathfrak{F}_\infty(\frac{1}{2}\Delta P)$ is self-adjoint it has no different self-adjoint extension. Therefore

$$\mathfrak{F}_\infty = \mathfrak{F}_\infty(\tfrac{1}{2}\Delta P).$$

In the same way we deduce from (13. 1)

$$\mathfrak{F}_\infty(\Delta P) = \mathfrak{F}_\infty(\tfrac{1}{2}\Delta P).$$

Thus Theorem 13. 1 is proved.

---

[27] Obviously they do not depend on c.

14. Example. Let the domain S consist of all points x except $x=0$; we take

$$r(x)=1$$

and define the form $\boldsymbol{P}$ by

$$v'\boldsymbol{P}v' = a(v_1^2 + \cdots + v_n^2) + 2b\left(\frac{x_1}{|x|}v_1 + \cdots + \frac{x_n}{|x|}v_n\right)v_0 + cv_0^2$$

where a, b, c are constants such that $a>0$, $c>0$, $ac-b^2>0$. Since the form $\boldsymbol{P}$ is positive-definite we may apply the theory of § 7 – § 12 and define spaces $\mathfrak{F}_\infty$ and $\mathfrak{F}_0$ in which the operator ED is self-adjoint.[28] This operator can be written as

$$ED = -a(D_1^2 + \cdots + D_n^2)u(x) - (n-1)\frac{b}{|x|}u(x) + cu(x).$$

On setting $n=3$, $a=\dfrac{\hbar}{2m}$, $b=\dfrac{e^2}{2}$, $c>\dfrac{me^4}{2\hbar}$, the operator $ED-c$ is Schrödinger's energy operator of the hydrogen atom. We, therefore, know that this operator is self-adjoint in the space $\mathfrak{F}$.

Usually, Schrödinger's operator or generally the above operator ED is connected with the quadratic form

$$v'\boldsymbol{P}_0v' = a(v_1^2 + \cdots + v_n^2) - (n-1)\frac{b}{|x|}v_0^2 + cv_0^2;$$

but this form is everywhere positive-definite only if $b\leqq 0$; then the difference form

$$v'(\boldsymbol{P}_0-\boldsymbol{P})v' = -2b\left(\frac{x_1}{|x|}v_1 + \cdots + \frac{x_n}{|x|}v_n\right)v_0 - (n-1)\frac{b}{|x|}v_0^2$$

is admissible in our sense. On the other hand, in case $b>0$ it is preferable to work with the form $\boldsymbol{P}$ instead of $\boldsymbol{P}_0$.

15. Continuity. In this paragraph we prove the continuity test mentioned in the introduction. To give it a precise form we introduce the space $\mathfrak{K}_r$ and operators Δ_r, which are applicable in $\mathfrak{K}_r$ and produce functions in $\mathfrak{H}$ if r is even, in $\mathfrak{H}'$ if r is odd. We define these spaces and operators by recursion.

---

[28] In case $n\geqq 2$ we have $\mathfrak{G}_\infty=\mathfrak{G}_0$; hence $\mathfrak{G}_0'=\mathfrak{G}'_\infty$ and $\mathfrak{F}_0=\mathfrak{F}_\infty$; thus both boundary conditions lead to the same space $\mathfrak{F}$. For the proof of this fact one may use the same reasoning as for Theorem 2.3 of my paper in the *Mathematische Annalen* Bd. 112, p. 8 (1935) (on setting $y=\displaystyle\int\frac{d\,|x|}{|x|^{n-1}}$).

1. $\mathfrak{K}_1 = \mathfrak{K}$; $\nabla_1 = D$.

2. $\mathfrak{K}_{2\rho}$ consists of all $u(x)$ in $\mathfrak{K}_{2\rho-1}$ for which $\nabla_{2\rho-1}u$ is in $\mathfrak{K}'$; $\nabla_{2\rho} = D^*\nabla_{2\rho-1}$.

3. $\mathfrak{K}_{2\rho+1}$ consists of all $u(x)$ in $\mathfrak{K}_{2\rho}$ for which $\nabla_{2\rho}u$ is in $\mathfrak{K}'$; $\nabla_{2\rho+1} = D\nabla_{2\rho}$.

We notice that the spaces $\mathfrak{K}_r$ actually do not depend on the number δ which appeared in the definition of D and D^* (cf. § 4).

Introducing the number

$$m = \left[\frac{n}{2}\right] + 1,$$

where n is the dimension of the x-space, we have

THEOREM 15. 1. *If $r \geqq m$ the functions $u(x)$ in $\mathfrak{K}_r$ are continuous and have continuous derivatives up to the order $r - m$.*[29]

Since the statement does not depend on the number δ we may choose

$$\delta = 0;\ \text{i.e.,}\ Du = \{0, D_1u, \cdots, D_nu\},\ D^*u' = - D_1u_1 - \cdots - D_nu_n.$$

We denote by ∇_r^* the formal adjoint of ∇_r; i.e., $\nabla_r^* = D^* \cdots D^*D$ for r even and $\nabla_r^* = D^* \cdots DD^*$ for r odd.

We introduce for even positive r the function

$$G^r(x) = |\, x\,|^{-(n-r)}(\gamma_r + \beta_r \log |\, x\,|)$$

and for odd positive r the system $G^r = \{0, G_1{}^r, \cdots, G_n{}^r\}$ where

$$G_\nu{}^r(x) = x_\nu \,|\, x\,|^{-(n-r+1)}(\gamma_r + \beta_r \log |\, x\,|),$$

defined for $x \neq 0$. The coefficients γ_r and β_r may be chosen such that the relations

$$\begin{aligned} DG^r &= - G^{r-1} \text{ for } r \text{ even,} \\ D^*G^r &= - G^{r-1} \text{ for } r \geqq 3 \text{ odd,} \\ D^*G^1 &= 0 \end{aligned}$$

hold. γ_r and β_r are uniquely determined by these relations, by the initial values $\gamma_1 = \Omega_n{}^{-1}$ where Ω_n is the area of the surface of the n-dimensional unit-sphere and by the condition $\gamma_n = 0$ for n even.

We notice that $\nabla_r^*G^r = 0$; thus G^r is the fundamental solution of the differential operator ∇_r^*.

---

[29] Cf. [8], [9], [10].

We introduce a function $\eta_R(t)$, which has derivatives of every order such that which $\eta_R(t) = 1$ for $|t| \leqq R$, $0 \leqq \eta_R(t) \leqq 1$ as $R \leqq t \leqq 2R$, $\eta_R(t) = 0$ as $|t| \geqq 2R$. Let $S' < S$ and a such that $S'_{2R} < S$ (cf. § 2). For x_0 in S' we set

$$k^r(x) = K^r(x_0, x) = \eta_R(x - x_0) G^r(x - x_0).$$

This function vanishes outside of S'_{2R} and

$$\nabla_r{}^* k^r(x) = \nabla_r{}^* K^r(x_0, x) = 0$$

there and in the neighborhood of $x = x_0$.

Let $v(x)$ be a function, defined in S, which has derivatives of every order. Then we can represent this function by means of the fundamental solution.

$$(15.1) \quad v(x_0) = -\int_S K^r(x_0, x) \cdot \nabla_r v(x) dx + \int_S \nabla_r{}^* K(x_0, x) v(x) dx.$$

We now use the operator J_a (cf. § 2) where the kernel function $j_a(x)$ has derivatives of every order. As an immediate consequence of Lemma 5.1 and 6.1 we have

$$\nabla_r J_a = J_a \nabla_r$$

and, in view of 2.2,

$$|\nabla_r J_a u - \nabla_r u|_{S'} \to 0 \text{ as } a \to 0$$

for every $u(x)$ in $\mathfrak{R}_r$.

We first prove that $u(x)$ in $\mathfrak{R}_r$ is continuous if $r \geqq m$. We take $r = m$. The functions $k^m(x)$ can be considered as functions in $\dot{\mathfrak{H}}$ as m is even, in $\dot{\mathfrak{H}}'$ for m odd; for

$$\int_{|x-x_0|<a} |k^m(x)|^2 dx = \Omega_n \int_0^a |x - x_0|^{2m-n-1} | \gamma_m + \beta_m \log |x - x_0| |^2 d |x - x_0| < \infty.$$

Let $u(x)$ be a function in $\mathfrak{R}_m$. We insert $v(x) = J_a u(x)$ in (15.1); clearly the right hand side converges uniformly for x_0 in S' to a continuous function $u^0(x_0)$. We have $|J_a u - u^0|_{S'} \to 0$ as well as $|J_a u - u|_{S'} \to 0$ (cf. 2.2); thus $|u - u^0|_{S'} = 0$; i. e., $u(x) = u^0(x)$ is in S' and, therefore, $u(x)$ is continuous.

In order to prove the statement concerning derivatives we introduce the n^{r-m} operators D^{r-m}

$$D^{r-m} = D_{\nu_1} D_{\nu_2} \cdot \cdot \cdot D_{\nu_{r-m}}.$$

We may apply D^{r-m} on both sides of (15.1) and obtain

$$(15.2)\quad D^{r-m}v(x_0) = -\int_S D^{r-m}_{(0)} K^r(x_0, x) \cdot \nabla_r v(x)\,dx + \int_S D^{r-m}_{(0)} \nabla_r{}^* K(x_0, x) \cdot v(x)\,dx$$

where the subscript (0) indicates that the operator acts with respect to x_0. The function $k(x) = D^{r-m}_{(0)} K^r(x_0, x)$ is in $\dot{\mathfrak{H}}$ as r is even, in $\dot{\mathfrak{H}}'$ as r is odd, since

$$|\, k(x)\,|^2 \,|\, x - x_0\,|^{n-1} \leqq \text{const}\,|\, x - x_0\,|^{2m-n-1}\,|\, \gamma_r + \beta_r \log |\, x - x_0\,|\,|^2$$

in the neighborhood of $x = x_0$.

Therefore we may apply the same reasoning as before and find that $D^{r-m}J_a u(x_0)$ tends to a continuous function $u^{r-m}(x_0)$ uniformly in S'. On assuming that our statement is true for $r - 1$ we know that $D^{r-m-1}J_a u(x_0)$ tends to $D^{r-m-1}u(x_0)$ uniformly in S'. Writing $D^{r-m}J_a u = D_r D^{r-m-1} J_a u$, we see that $D^{r-m-1}u(x)$ has the continuous derivative $u^{r-m}(x)$. Thus theorem 15.1 is proved.

We introduce a space $\mathfrak{R}^s$ and n^s operators D^s in $\mathfrak{R}^s$ by the recursion: $\mathfrak{R}^1 = \mathfrak{R}$ and $D^1 = D_\nu$, $\nu = 1, \cdots, n$; $\mathfrak{R}^s$ consists of all $u(x)$ in K^{s-1} for which each $D^{s-1}u(x)$ is in $\mathfrak{R}$; $D^s u(x) = D_\nu D^{s-1} u(x)$, $\nu = 1, \cdots, n$. Then we state

THEOREM 15.2. *Let $u(x)$ be in $\mathfrak{R}_{2\rho}$, $\nabla_{2\rho}u(x)$ be in $\mathfrak{R}^s$. Then $u(x)$ is in $\mathfrak{R}^{s+1}$, each $D^s u(x)$ is in $\mathfrak{R}_{2\rho}$ and*

$$D^s \nabla_{2\rho} u = \nabla_{2\rho} D^s u = \nabla_{2\rho-2} D^* D D^s u.$$

It is sufficient to prove the statement for $s = 1$; for then the statement follows by induction for $s > 1$.

From (15.1) we get by differentiation and partial integration

$$D_\nu v(x_0) = -\int_{S'_{2R}} K^{2\rho}(x_0, x) D_\nu \nabla_{2\rho} v(x)\,dx - \int_{S'_{2R}} D_\nu \nabla^*{}_{2\rho} K^{2\rho}(x_0, x) v(x)\,dx,$$

observing that the integrands vanish outside of S'_{2R}.

From this we obtain the inequality

$$|\, Dv\,|_{S'} \leqq \int_{|x|\leqq 2R} |\, K^{2\rho}(x)\,|\,dx\,|\, D\nabla_{2\rho} v\,|_{S'_{2R}} + \int_{|x|\leqq 2R} |\, D\nabla^*{}_{2\rho} K^{2\rho}(x)\,|^2 dx\,|\, v\,|_{S'_{2R}}$$

by the same reasoning that led to (2.1), taking $K^{2\rho}(x)$ and $D_\nu \nabla^*{}_{2\rho} K^{2\rho}(x)$ instead of $j_a(x)$. We now set $v(x) = J_a u(x)$; we have $|\, J_a u - u\,|_{S'_{2R}} \to 0$

and, since $D\nabla_{2\rho}J_a u(x) = \nabla_{2\rho}u(x)$, also $| D\nabla_{2\rho}J_a u - D\nabla_{2\rho}u |_{S'_{2R}} \to 0$. Hence $| DJ_a u - DJ_b u |_{S'} \to 0$ as $a, b \to 0$. Therefore a system $u'(x)$ in $\mathfrak{L}'$ exists such that $| DJ_a u - u' |_{S'} \to 0$. This means, according to 5.3, that $u(x)$ is in $\mathfrak{K}$ and $Du = u'$.

We apply the same reasoning to $\nabla_r D_\nu v(x_0)$, taking $(-1)^r \nabla_r K^{2\rho}(x_0, x)$ instead of $K^{2\rho}(x_0, x)$; we obtain that $D_\nu u$ is in $\mathfrak{K}_r$, $r = 1, \cdots, 2\rho$ and we deduce $\nabla_{2\rho} D_\nu u = D_\nu \nabla_{2\rho}$ from $\nabla_{2\rho} D_\nu J_a u = D_\nu \nabla_{2\rho} J_a u$. Thus Theorem 15. 2 is proved.

In immediate consequence of Theorem 15. 2 is

THEOREM 15. 3. *Let $q(x)$ and $f(x)$ be functions defined in S where $q(x)$ has continuous derivatives up to the order s while $f(x)$ is in $\mathfrak{K}^s$. If u in $\mathfrak{K}_2$ is a solution of*

$$D^*Du + qu = f$$

then u is in $\mathfrak{K}^s$; if $s \geqq m$ then u is continuous and has continuous derivatives up to the order $s - m$.

First we observe that qu is in $\mathfrak{K}^\sigma$ if u is in $\mathfrak{K}^\sigma$; $\sigma = 1, \cdots, s$. For $\sigma = 1$, this is an immediate consequence of 5. 2 applied on qu and $qDu + Dqu$; for $\sigma > 1$ we may apply induction.

Since u is in $\mathfrak{K}^1$, also $D^*Du = f - qu$ is in $\mathfrak{K}^1$; hence we learn from Theorem 15. 2 that u is in $\mathfrak{K}^2$; therefore $D^*Du = f - qu$ is in $\mathfrak{K}^2$, too, and u is in $\mathfrak{K}^3$.

Repeating this reasoning we find that u is in $\mathfrak{K}^s$. Since, obviously, $\mathfrak{K}^s$ contains $\mathfrak{K}_s$ we may apply Theorem 15. 1 in case $s \geqq m$.

In view of the remarks made in the introduction this theorem implies that the characteristic functions of the operator $D^*D + q$ have continuous derivatives up to the order $s - m$ if q has continuous derivatives up to the order s.

NEW YORK UNIVERSITY,
NEW YORK, N. Y.